**References**

1 Keller, M. and Zengler, K. (2004) Tapping into microbial diversity. *Nat. Rev. Microbiol.* 2, 141–150
2 DeLuca, T.H. *et al.* (2002) Quantifying nitrogen-fixation in feather moss carpets of boreal forests. *Nature* 419, 917–920
3 Iverson, V. *et al.* (2012) Untangling genomes from metagenomes: revealing an uncultured class of marine Euryarchaeota. *Science* 335, 587–590
4 Helbling, D.E. *et al.* (2012) The activity level of a microbial community function can be predicted from its metatranscriptome. *ISME J.* 6, 902–904
5 Cardinale, B.J. *et al.* (2006) Effects of biodiversity on the functioning of trophic groups and ecosystems. *Nature* 443, 989–992
6 Tylianakis, J.M. *et al.* (2008) Global change and species interactions in terrestrial ecosystems. *Ecol. Lett.* 11, 1351–1363
7 Zhou, J. *et al.* (2012) Microbial mediation of carbon-cycle feedbacks to climate warming. *Nat. Clim. Change* 2, 106–110
8 Tyson, G.W. *et al.* (2004) Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* 428, 37–43
9 Woyke, T. *et al.* (2006) Symbiosis insights through metagenomic analysis of a microbial consortium. *Nature* 443, 950–955
10 Stein, L.D. (2010) The case for cloud computing in genome informatics. *Genome Biol.* 11, 207

# Let the four freedoms paradigm apply to ecology

## Duccio Rocchini and Markus Neteler

Fondazione Edmund Mach, Research and Innovation Centre, Department of Biodiversity and Molecular Ecology, Via E. Mach 1, 38010 S. Michele all'Adige (TN), Italy

In 1985, Richard Stallman, one of the most brilliant minds in computer science, founded the Free Software Foundation and launched the concept of 'copyleft', the opposite of copyright. The aim, outlined in the GNU Manifesto (http://www.gnu.org/gnu/manifesto.html, [1]), was to make software programs 'free' as in 'freedom'.

The famous 'four freedoms' expounded by Stallman [1] are: (i) the freedom to run the program for any purpose; (ii) the freedom to study how the program works and adapt it to one's own needs; (iii) the freedom to redistribute copies; and (iv) the freedom to make improvements to the program and release them to the public. Thus, the whole (scientific) community benefits from software development. These freedoms are also inherent in several free software licenses, the GNU General Public License (GPL) being one of the most popular.

Approximately a quarter of a century after Stallman put forward his ideas, William K. Michener and Matthew B. Jones, in an article in *TREE* [2] focusing on the analysis of ecological data, stated that: 'analytical processes are fundamental to most published results in ecology'. Explicit reference to the analytical procedures adopted in generating scientific results is crucial for reproducibility, yet these processes are rarely documented in published ecological papers [2]. Scientific workflow applications, such as Kepler (https://kepler-project.org), attempt to address the problem [2], but are only partially successful because the underlying algorithms may still be opaque.

In our view, the explicit use of Free and Open Source Software (FOSS) with availability of the code is essential for completely open science: 'scientific communication relies on evidence that cannot be entirely included in publications', but 'anything less than the release of source programs is intolerable for results that depend on computation' [3].

The idea of FOSS and the public availability of the code has been around for almost as long as software [4]. Nonetheless, as far as ecologists are concerned, the open source philosophy is only just taking off, as Stokstad has also pointed out [5].

The increasing availability of open ecological data through networks such as the Global Biodiversity Information Facility (GBIF, http://www.gbif.org, [6]) or the Data Observation Network for Earth (DataONE) federated data archive (http://www.dataone.org, [7]) makes it increasingly possible to test cutting-edge ecological theories, such as dark diversity [8], evolutionary paths [9] and climate change scenarios [10]. In using a shared open-source code for testing these ecological theories, researchers can be sure that their results are reliable and also that the code they have used is robust [11]. This is particularly true when complex algorithms (or statistical approaches) are involved.

To avoid black box calculations and built-in user interfaces, criticized in [2], researchers have recourse to several examples of FOSS in areas of ecological research, such as ecological statistics (e.g. R Language and Environment for Statistical Computing, http://www.R-project.org, [12]) and spatial ecology [e.g. Geographical Resources Analysis Support System (GRASS) GIS, http://grass.osgeo.org, [4]). The modular design of such software means decentralized contributions can be made to the source code and allows different institutions and individuals around the world to improve the code base.

If FOSS were more widely employed in ecology and the code used in data analysis provided in scientific papers, more researchers [11] would be able to rely on and replicate

*Corresponding author:* Rocchini, D. (ducciorocchini@gmail.com), (duccio.rocchini@fmach.it).

peer-reviewed functions. Efforts still need to be made in this area to improve the processes for sharing what is in effect the backbone of ecological software: its code. Therefore, there is an urgent need to embrace Stallman's four freedoms paradigm in ecology.

### References

1 Stallman, R. (1985) The GNU Manifesto. *Dr Dobb's J.* 10, 30
2 Michener, W.K. and Jones, M.B. (2012) Ecoinformatics: supporting ecology as a data-intensive science. *Trends Ecol. Evol.* 27, 85–93
3 Ince, D.C. *et al.* (2010) The case for open computer programs. *Nature* 482, 485–488
4 Neteler, M. *et al.* (2012) GRASS GIS: a multi-purpose Open Source GIS. *Environ. Model. Softw.* 31, 124–130
5 Stokstad, E. (2011) Open-Source ecology takes root across the world. *Science* 334, 308–309
6 Guralnick, R.P. *et al.* (2007) Towards a collaborative, global infrastructure for biodiversity assessment. *Ecol. Lett.* 10, 663–672
7 Reichman, O.J. *et al.* (2011) Challenges and opportunities of open data in ecology. *Science* 331, 703–705
8 Pärtel, M. *et al.* (2011) Dark diversity: shedding light on absent species. *Trends Ecol. Evol.* 26, 124–128
9 Scally, A. *et al.* (2012) Insights into hominid evolution from the gorilla genome sequence. *Nature* 483, 169–175
10 Moss, R.H. (2010) The next generation of scenarios for climate change research and assessment. *Nature* 463, 747–756
11 Barnes, N. (2010) Publish your computer code: it is good enough. *Nature* 467, 753
12 R Development Core Team (2012) *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, (Vienna, Austria)

# Type I error hinders recycling: a response to Rohr and Martin

## Daniel Montesinos

Centro de Ecologia Funcional, Universidade de Coimbra, Apartado 3046, 3001-401, Coimbra, Portugal

Rohr and Martin [1] propose that scientific reviews should be recycled in order to reduce the burden of repeated peer-review processes. Many other interesting initiatives are already reducing such burdens in different ways. For instance, and to cite only two examples, PLoS ONE is publishing papers based strictly on scientific quality and not on competition for limited space in a printed journal. This policy naturally reduces the number of reviews that a paper needs before publication, since reviews are only aimed to select *good* (and not *better*) manuscripts. A different enterprise, Peerage of Science, is trying to externalize the review process from journals. Although the system viability raises some doubts, there are also clear advantages [2,3]. Under the Peerage of Science model, journals pick papers that have had a good set of reviews and fit the journal scope and editorial policy. Since there is only one round of reviews, the reduction in the number of reviews needed to publish is evident. However, there is a risk that some papers might never be picked by any journal. The ideas of Rohr and Martin come along with all these novelties and are worth considering; however, their model presents some apparent problems.

The risk of having a Type I error publication bias (false positive) is inherent to any peer review process. At least theoretically, a bad scientific manuscript, if submitted a sufficient number of times, should eventually get published 'just by chance' (a.k.a. 'good luck with the referees'). Although we all want to believe that the number of submissions needed to publish a bad manuscript just by chance is larger than the patience of a scientist, chances are that it might happen now and then (up to five out of 100 times, according to our current statistical dogma). Therefore, in a perfect editorial world, we should probably need a Bonferroni-like correction (or a Bayesian prior), in which a manuscript reviewed several times will need an increasing number of positive reviews in order to get published.

In the system proposed by Rohr and Martin, if an author decides *when* to forward a review, and decides *which* reviews are to be forwarded, there is a clear risk that only positive reviews will be forwarded, and that just by accumulation of those positives the manuscript will eventually get published. Worryingly, the number of submissions needed before an editor commits a Type I error might become dangerously low. If on each review process only half of the referees provide reasonably positive reviews, after only two or three submissions a questionable manuscript could be accepted for publication. An author might even want to be rejected by some mid-tier journals and then forward a selected collection of positive reviews to a top-tier journal and get their article published.

Some mechanisms could be developed to try to avoid these undesirable problems. For instance, editors might be willing to forward to other editors a *full set* of reviews, at the authors' request. Although this will alleviate the abovementioned problem, there is still a chance that more positive sets of reviews will be requested to be forwarded more often than more negative ones, resulting again in a milder but still significant risk of committing Type I errors. Besides, how many authors would wish to be evaluated

*Corresponding author:* Montesinos, D. (danimontesinos@gmail.com).