164 Ashdale Avenue
Toronto, Ontario
Canada M4L 2Y9
+1 (416) 435 9779
gvwilson@software-carpentry.org
http://software-carpentry.org
June 3, 2013

To Whom It May Concern,

Re: Pre-Submission Inquiry for “Best Practices in Scientific Computing”

We are writing to ask whether *PLoS Biology* would be interested in publishing our paper, “Best Practices for Scientific Computing”. Our motivation for writing it is the realization that while computing is now central to all forms of scientific research, few scientists are aware of good practices for developing and using software rigorously and efficiently [1,2]. As a result, their work is often not reproducible and can include serious errors that can mislead the development of scientific ideas for years (e.g, [3,4,5]). Just as importantly, the lack of good “lab skills” for computational work means that many researchers spend far more time than necessary doing simple things. A growing recognition of these problems has triggered a flurry of editorials, commentary pieces, and reviews (e.g., [6,7,8]). However, none of these have explained what scientists should actually do to improve the reproducibility, correctness, and efficiency of their work.

Our paper is intended to fill that gap. In 4800 words, we describe 27 practices that can improve the quality and productivity of computational work, and provide citations to empirical software engineering research to support our case (e.g., [9-12]). Rather than providing a “one size fits all” prescription, we base our recommendations on the experience we have gained through the Software Carpentry project, and specifically from running almost a hundred two-day “boot camps” for over 2500 scientists since January 2012.

A preprint of this paper has been cited half a dozen times since being posted to arXiv in October 2012 [13], and is one of the starting points for a two-day NIH-sponsored workshop this summer discussing what biologists should be taught about computing. We believe this shows that the paper is timely, and likely to be of interest to your readership. We hope you will agree, and look forward to your response.

Sincerely,

Dr. Gregory V. Wilson

Mozilla Foundation / Software Carpentry

[1] J.E. Hannay et al.: “How Do Scientists Develop and Use Scientific Software?” *Proc. 4th International Workshop on Software Engineering for Computational Science and Engineering*, 2009; DOI:10.1109/SECSE.2009.5069155.

[2] P. Prabhu et al.: “A Survey of the Practice of Computational Science.” *Proc. 24th ACM/IEEE Conference on High Performance Computing, Networking, Storage and Analysis*, 2011; doi.acm.org/10.1145/2063348.2063374.

[3] G. Chang: “Retraction of 'Structure of MsbA from *Vibrio cholera*: A Multidrug Resistance ABC Transporter Homolog in a Closed Conformation' [J. Mol. Biol. (2003) 330 419–430]” *Journal of Molecular Biology*, 369(2), 2007.

[4] D.C. Lees and R.K. Colwell: “A strong Madagascan rainforest MDE and no equatorward increase in species richness: re-analysis of 'The missing Madagascan mid-domain effect', by J.T. Kerr, M. Perring & D.J. Currie [Ecology Letters 9:149–159, 2006]” *Ecology Letters*, 10(9), 2007.

[5] D.A. Kelt et al.: “Differential Responses of Two Species of Kangaroo Rat (*Dipodomys*) to Heavy Rains: A Humbling Reappraisal.” *Journal of Mammalogy*, 89(1), 2008.

[6] A. Morin et al.: “Shining Light into Black Boxes.” *Science*, 2012; DOI:10.1126/science.1218263, http://www.sciencemag.org/content/336/6078/159.summary

[7] Z. Merali: “Computational Science: …Error” *Nature*, 2010; DOI:10.1038/467775a, http://www.nature.com/news/2010/101013/full/467775a.html

[8] G. Miller: “A Scientist’s Nightmare.” *Science*, 2006; DOI:10.1126/science.314.5807.1856, http://www.sciencemag.org/content/314/5807/1856.summary

[9] J. Carver et al.: “Software Development Environments for Scientific and Engineering Software: A Series of Case Studies.” *Proc. ICSE 2007.*

[10] D. Matthews et al.: “Configuration Management for Large-Scale Scientific Computing at the UK Met Office.” *Computing in Science & Engineering*, Nov-Dec 2008.

[11] A. Oram and G. Wilson (eds.): *Making Software.* O'Reilly, 2010, 0596808321.

[12] J. Segal: “When Software Engineers Met Research Scientists: A Case Study.” *Empirical Software Engineering*, 10(4), 2005.

[13] G. Wilson et al.: “Best Practices in Scientific Computing.” http://arxiv.org/abs/1210.0530.