

Comment: the case for open preprints in biology

Martin Fenner, Gobbledygook

March 30, 2013

Last week Philippe Desjardins-Prouly et al. published the article The case for open preprints in biology – naturally as a preprint on figshare. The article sees preprint servers as a great opportunity for open science, and discusses the status of preprints in the biological sciences. In this blog post I want to add some comments to the text.

E-BIOMED

What is now PubMed Central started out as E-BIOMED in 1999 and initially was envisioned to include a repository for preprints. It is important to look back at what happened then, and why the preprint repository was dropped from what then became PubMed Central. Harold Varmus talks a bit about this in this interview from 2006.

Nature Precedings

The article talks about why biologists have not developed a culture of sharing preprints. It would be good to mention Nature Precedings, a preprint server for the life sciences started in 2007 that stopped taking new submissions in 2012. This blog post on RetractionWatch cites the announcement by Nature Publishing Group (which doesn't explain why the service was shut down), and there are a good number of interesting comments.

SSRN

Preprints in other disciplines are mentioned in the text, in particular ArXiv, but also RePEc. I would also include **SSRN** (Social Science Research Network), which uses a different model, but is as important for the working paper and preprint culture in the social sciences as ArXiv is in physics/mathematics.

Google Scholar Metrics

In April 2012 Google launched Google Scholar Metrics, listing the top 100 publications (according to their h5-index) in several disciplines. Six out of the top 10 publications in physics/mathematics are ArXiv sections (arXiv

Astrophysics (astro-ph) is #2), the IZA Discussion Papers are #1 in Social Sciences, and the NBER Working Papers are #1 in Economics, and arXiv Astrophysics (astro-ph) is #12 on the top 100 list for all disciplines (#1-5 are journals in biology and medicine: *Nature*, *New England Journal of Medicine*, *Science*, *Lancet*, *Cell*). All these metrics are a strong indicator that preprints can be highly cited.

Citation Advantage of Preprints

Anne Gentil-Beccot et al. have written a nice paper (of course available as preprint) that shows that publication as preprint now only increases the citation rate for the corresponding peer-reviewed article published later, but also leads to much faster citations, with a peak immediately after publication.

Average number of citations per article per month as a function of the time of the citation relative to the time of publication. From <http://arxiv.org/abs/0906.5418>

SCOAP3

The Sponsoring Consortium for Open Access Publications in High Energy Physics (SCOAP3) is working on turning the majority of peer-reviewed publications in high energy physics into gold open access. It is important to understand that the high energy physics community feels that they need peer-reviewed journal articles in addition to ArXiv.

Preprint Culture in Clinical Medicine

It is a little known fact that there is a strong preprint culture in clinical medicine. I have written about this topic in October 2010. Clinical trials have to be registered before starting the trial, and information about the trial is publicly available in clinicaltrials.gov and other registries. Results are presented in conferences (as poster or oral presentation), at which stage it becomes public information. The peer-reviewed paper – with a few exceptions – follows much later, sometimes even after drug approval by the FDA (in the blog post I used the TROPIC trial as example). The problem is of course that information in oral presentations and posters is incomplete and difficult to find. But publication of a clinical trial in a peer-reviewed journal is more about giving credit to the researchers involved (similar to SCOAP3 in high energy physics) than about spreading the knowledge. Peer review is not an appropriate filter for whether or not a new drug or drug combination should be used to treat patients – the approval process by regulatory authorities is much more extensive than any peer review can ever be.

Preprint culture in Biology

The paper mentions several reasons why the field of biology has essentially no preprint culture. One argument against preprints is that it would be easier to

steal ideas. Although I agree with the authors that preprints are a great way to establish precedence, there is a big difference between research based on years of work using expensive equipment (as is often the case in high energy physics but also some other fields), and research that can be reproduced in a few weeks. In the latter case it is possible that someone else is faster in publishing the peer-reviewed paper. Another difference is the community: “stealing” ideas from someone else is probably more difficult in smaller scientific communities, and some scientific communities are more competitive and less collaborative than others.

Another concern about preprints raised in the paper is the Ingelfinger rule, i.e. the uncertainty that a journal would accept a manuscript if already published as a preprint. This concern is fortunately unfounded regarding most publishers, and the paper includes a table listing the preprint policies of important publishers in biology.

I would like to add two other reasons why the preprint culture is probably not established in biology. Preprints are competition for the peer-reviewed journal article and scholarly publishers might not be particularly interested in encouraging a preprint culture. A lot has fortunately changed since E-BIOMED in 1999.

Finally, whereas some disciplines use preprints and working papers to communicate, in biology the preferred way to communicate research findings before publication of a peer-reviewed paper is the oral presentation. What we may need is a service that makes it easy to upload and share scientific presentations. We for example already have Slideshare, Speaker Deck as generic tools, and SciVee, fishare aimed at scientists. Speaker Deck is currently my favorite tool and is a Github product (Github has been mentioned in the manuscript as an option for hosting preprints). Maybe what is missing is a killer combination of features in a new or existing service – persistent identifiers, uploading of background material (text, data, software, video) in addition to the slides, non-textual search, cooperation with conference organizers, etc. – for presentation sharing to take off as a way to establish a preprint culture in biology.

Update 4/4/13: Yesterday PeerJ launched a new preprint service for life sciences research. Read this blog post for details, and this post on the Mendeley blog.

References

Philippe Desjardins-Proulx, Ethan P. White, Joel Adamson, Karthik Ram, Timothée Poisot, Dominique Gravel. The case for open preprints in biology. Figshare; 2013. Available from: <http://dx.doi.org/10.6084/M9.FIGSHARE.655710>