

preprints

Patrick D. Schloss[†]

[†] To whom correspondence should be addressed: pschloss@umich.edu

Department of Microbiology and Immunology, University of Michigan, Ann Arbor, MI

1 **Abstract**

2 **Importance**

3 PeerJ Preprints F1000 Zenodo / FigShare / GitHub

4 Thoughts. . . * Posters and talks at scientific conferences * Data are often deposited without any
5 expectation of peer review

6 Many scientists, including microbiologists, have begun to use preprints and other online venues
7 such as social media, blog posts, and videos as methods to engage their peers and non-scientists.
8 Preprints have emerged as a formal venue for microbiologists to disseminate their research. A
9 preprint is an unpublished manuscript that is made publicly available without going through an
10 official peer-review process. Rather, an author can post their manuscript to a preprint server for
11 others to read, share, and make comments. Preprints were initially adopted among physicists in the
12 1960s as a method of sharing interesting research amongst colleagues. Eventually, this community
13 evolved into what is now the arXiv (pronounced “archive”) preprint server that was hosted at the Los
14 Alamos National Laboratories from 1991 to 1999 and then at Cornell University. Among physicists
15 and mathematicians, posting a preprint to arXiv has become the standard publication strategy
16 followed by submission to a peer-reviewed journal. Within those communities researchers have
17 been recognized with international awards for work that has only been posted to arXiv without
18 formal review. Although arXiv has hosted a number of computational biology papers, the server
19 has not drawn widespread attention from biologists. Among the proponents of arXiv, preprints have
20 aided in the development of research communication by accelerating the release of the science
21 and helping it to achieve a wider audience for critique and reception. As adoption of preprints as
22 a component of the scientific method has broadened among biologists, we sought to explore the
23 specific use and concerns of preprints amongst microbiologists.

24 ***Acceptance of preprints by journals.*** An early controversy encountered by researchers
25 interested in pursuing preprints as a stage in disseminating their research was whether it
26 constituted prior publication. The broad consensus at this point is that preprints do not constitute
27 prior publication. To briefly summarize journals of interest to many microbiologists, journals
28 published by ASM, the Microbiology Society, International Society for Microbial Ecology, PLOS, the
29 Proceedings of the National Academy of Science, Science, and Nature have a permissive stance
30 towards prior posting of preprints prior to submission. Although journals published by Cell do not
31 forbid authors from posting preprints prior to submission, they ask authors to consult an editor prior
32 to posting and do not allow authors to post revised preprints that contain revisions that respond to
33 editorial input. Considering the relatively fluid nature of many of these policies and the journals’
34 specific policies, prospective authors should be aware of the positions taken by the journals where

they may eventually submit their work. Comprehensive lists of journals' attitudes towards preprints are available online and are regularly updated.

Landscape of preprint servers. Two preprint servers have sought to parallel the success of arXiv. In 2013, the bioRxiv (pronounced "bio-archive") was launched through Cold Spring Harbor Laboratory as a service to complement arXiv; PeerJ Preprints was initiated at the same time. Both options have similar features: preprint posting is free; each preprint receives a digital object identifier (DOI) that facilitates the ability to cite preprints in other scholarly work; if the preprint is ever published, the preprint is linked to the published version; the submission process for both options is relatively simple allowing authors to upload a PDF version of their preprint and supplemental materials; preprints are typically posted in under 24 hours; they have built in venues for authors to discuss their research with people who leave comments on the manuscript; preprints undergo a basic screening process to remove submissions with offensive or non-scientific content; and the sites provide article-level metrics indicating the number of times an abstract has been accessed or the preprint has been downloaded. There are several important differences between the two options. First, PeerJ as a for-profit organization and bioRxiv is a non-profit organization. This can be meaningful to authors since some journals, including the ASM Journals, will only accept submissions that have been posted on preprint servers hosted by non-profit organizations. Second, preprints at PeerJ Preprints are posted under the Creative Commons Attribution License (CC-BY) and bioRxiv preprints can be posted under one of four CC-BY licenses or with no permission for reuse. This can be relevant for authors hoping to submit their work a journal as journals will not consider manuscripts posted as preprints under a CC-BY license (e.g. *Proceedings of the National Academy of Sciences*). A cosmetic, but important difference between the two is the layout and feel of the two websites. Compared to the bioRxiv site, the PeerJ Preprint site is more fluid, gives readers the ability to "follow" a preprint, and provides better access to article keywords and the ability to search preprints. With broader acceptance of preprints by traditional journals, many journals, including all of the ASM journals, have established mechanisms to directly submit manuscripts that are posted as preprints on bioRxiv. It is only possible to transfer a PeerJ Preprint for submission to PeerJ. In many ways, preprint servers have taken on the feel of a journal. As adoption of this approach to disseminating research expands, it is likely that the features of these

sites will continue to improve. It is also likely that interfaces from third-parties will improve. For example, although Google Scholar includes preprints hosted at bioRxiv and PeerJ Preprints in their search results, PubMed does not. There is hope that the National Institutes of Health (NIH) will renew their interest in including preprints in PubMed search results.

Preprints and peer-review. The traditional peer-review editorial process was adapted to the technologies and trends that have evolved over the past 100 years. The formal peer-review system that most journals currently use was not developed until the end of the 1800s with the advent of typewriters and carbon paper (doi:10.1016/S0167-7799(02)01985-6). Editorial decisions were typically made by a single person or a committee (i.e. the editorial board) who had an expertise that covered the scope of the journal. As people's science became more specialized new journals would form providing a source of validation to the new specialty. The growth in science in the mid 1900s resulted in a shift from difficulties finding sufficient numbers of manuscript to publish to having too many manuscripts submitted. It has been argued that the widespread adoption of peer review was due to the increased specialization and to deal with the large number of manuscript submissions (JAMA 1990;263:1323-1329). Peer-review did not achieve widespread use at many journals, including the Journal of Bacteriology, until the 1940 and 1950s. Given the rapid advances in communication technology and even greater specialization, it is worth pondering whether the current scientific publishing system and peer review system, in particular, needs to continue to adapt with our science.

Communicating research has traditionally be done within research group meetings, departmental seminars, conferences, and as publications. Along this continuum, there is the assumption that the quality of the science has been improved because it has been vetted by more experts in the field. The public dissemination of one's research is a critical component of the scientific method. By describing their research, scientists subject their work to formal and informal peer review. Their research is scrutinized, praised, and probed to identify questions that help seed the next iteration of the scientific method. A common critique of these more modern approaches to publishing has been an inability to assess the quality of the science without the validation of peer-review. Attached to assertions of the validity of the research has become assertions of the impact and robustness of the research. These are all quality assessments that many acknowledge are difficult to assess

by the traditional peer-review process. This has led to some to call for peer reviewers to place a reduced emphasis on the impact or significance of the work as pioneered by the journal PLOS ONE. It has also led to the call for replacing or complementing pre-publication peer review with post-publication peer review using PubMed Commons, PubPeer, journal-based discussion forums, and other mechanisms. Alas if scientists are going to depend on post-publication peer review or informal methods of peer-review, they must be willing to provide constructive feedback on the work of others.

Preprints have the potential to change how science is measured. Preprints are often viewed as existing in a state of scientific limbo. They represent a formal communication, but are not “published”. As the use of preprints grows and our perceptions of preprints matures, there are a number of issues that will need to be addressed. First, a common concern is that if a researcher posts their work as a preprint, it will be “scooped” by another researcher and the original researcher will lose their ability to claim primacy or their ability to publish the work in a journal. Considering the preprint is a citable work with a DOI, it would, in fact, be the original researcher that scooped the second. A growing number of scientific societies and journals, including ASM view preprints as citable and as a legitimate claim to primacy. Some worry that with such protection a researcher can make a claim without valid data to back it up. This is possible; however, it is also the responsibility of the scientific community to utilize the peer-review mechanisms that are available to comment on those preprints pointing out methodological problems or to indicate that they are speaking beyond the data. A second area of concern is whether a preprint can be used to support a grant proposal. Given the length limitations of many funding agencies, there is a push to cite previous work to indicate a research team’s competence in an area or to provide preliminary data. Some fear that the use of preprints will allow some to circumvent page limits by posting a half-baked manuscript. We would hope that both consumers of preprints and grant proposal reviewers would be able to differentiate between someone trying to game the system and someone that is using preprints as a mechanism to improve their science. A third concern is what role preprints should have in assessing a scientist’s productivity. Clearly use of publication metrics is a contentious topic without considering the place of preprints. Regardless, given the propensity for researchers to list manuscripts as being “in preparation” on an application or curriculum vitae, listing them instead

as preprints that can be reviewed by a committee could be seen as a significant way to enhance an application. In fact, several funding agencies are starting to encourage fellowship applicants to include preprints in their materials and others are requiring researchers to post preprints for all of their work. It is clear that the adoption of preprints will challenge how scientists interact and evaluate each other's work; however, it appears that there is great potential to empower researchers by controlling when a citable piece of work is made public.

The use of preprints from an author's perspective. Posting research as a preprint gives an author great control over when their work is made public. Under the traditional peer-review model, an author may need to submit and revise their work multiple times over a long period. In contrast, an author can post the preprint at the start of the process for others to consume and comment on as it works its way through the editorial process. A first example illustrates the utility of preprints for improving access to research and the quality of its reporting. In 2014, the Schloss laboratory posted a preprint to PeerJ Preprints describing a method of sequencing 16S rRNA gene sequences using the Pacific Biosciences sequencing platform. At the same time, we submitted for review at PeerJ. While it was under review, we received feedback from someone that read the preprint as well as from Pacific Biosciences that the impact of the results could be limited because of the sequencing chemistry. Instead of resubmitting the manuscript to address the reviews, we generated new data and submitted an updated preprint a year later with a simultaneous submission to PeerJ that incorporated the original reviews as well as the feedback we received from Pacific Biosciences and the independent scientist. In general, the Schloss laboratory has simultaneously posted our manuscripts as preprints and submitted them to a journal. As a second example, this manuscript was posted to bioRxiv as a preprint on XXXXXXXX XX, 2017. We then solicited feedback on the manuscript using social media. Two weeks later, we incorporated the comments and posted a revised preprint and submitted the manuscript to *mBio*. During those two weeks the abstract was read XXXX times and the PDF was accessed XXXX times. This process engaged XXXX commenters on bioRxiv, XXXX people on Twitter, XXXX on Facebook, and XXXX via email. We received useful feedback from XXX people. Compared to the two or three scientists that typically review a manuscript, this experience engaged a much larger and more diverse community than had we foregone the posting of a preprint. Although there are concerns regarding the quality of the

science posted to a preprint server, we contend that responsible use of preprints as a part of the scientific process can significantly enhance the science.

Preprints in microbiology

- Rate of use among PP describing microbiology research over time
- Number of revisions
- Number of downloads
- Number of PP published

Microbiology anecdotes

- Minlon methods
- PacBio paper
- Tardigrade microbiome
- NYC subway microbiome

Specific challenges for microbiology

- Dual use
- Protection of human subjects
- Conflicts of interest

Perspective

- Support priority in literature
- Support citation
- Direct submission
- preprint scouts

- 172 • mSphereDirect / F1000Prime
- 173 • Licensing
- 174 • What's in it for ASM Journals and Readers?

