The early bird uses bioRxiv: The impact of career stage on the usage of preprints in ecology and evolution

J.F. Wolf[1] (he/him) (https://orcid.org/0000-0003-0773-4456), L. MacKay[2,3] (she/her), S.E. Haworth[1,3] (she/her), M.L. Cossette[1,3] (she/her), M.N. Dedato[1,3] (she/her), K.B. Young[1,3] (she/her), C.I. Elliott[2,3] (he/him), R.A. Oomen[4,5] (she/her) (https://orcid.org/0000-0002-2094-5592)

[1] Trent University, Department of Environmental and Life Sciences, Peterborough, Ontario, K9L 0G2, Canada

[2] Trent University, Department of Forensic Science, Peterborough, Ontario, K9L 0G2, Canada

[3] These authors contributed equally: L. MacKay, S.E. Haworth, M.L. Cossette, M.N. Dedato, K.B. Young, C.I. Elliott

[4] Centre for Ecological and Evolutionary Synthesis, Department of Biosciences, University of Oslo, 0371 Oslo, Norway

[5] Centre for Coastal Research, Department of Natural Sciences, University of Agder, 4604 Kristiansand, Norway

Corresponding author: Rebekah A. Oomen, rebekahoomen@gmail.com, Centre for Ecological and Evolutionary Synthesis, Department of Biosciences, University of Oslo, 0371 Oslo, Norway

Keywords

Authorea, Early Career Researcher, ECR, EcoEvoArxiv, preprint servers

Abstract

The usage of preprint servers in ecology and evolution is increasing, as it allows for research to be rapidly disseminated and available through open access at no cost. This is relevant for Early Career Researchers (ECRs), who must demonstrate research ability for funding opportunities, scholarships, grants, or faculty positions in short temporal windows in order to advance their careers. Concurrently, limited experience with the peer review process can make it challenging for those who are in the early stages of their research career to build publication records. Therefore, ECRs face different challenges relative to researchers with permanent positions and established research programs and have different requirements in terms of research output and timelines. These challenges might also vary according to institution size and country, which are associated with the availability of funding for open access journals. Herein, we hypothesize that career stage and institution size impact relative usage of preprint servers among researchers in ecology and evolution. Using data collected from 500 articles (100 from each of two open access journals, two closed access journals, and a preprint server), we demonstrate that ECRs generate more preprints relative to non-ECRs, for both first and last authors. We speculate that this pattern is reflective of the advantages of quick and open access research that is disproportionately beneficial to ECRs. There is also a marginal effect of first author institution size on preprint usage, whereby the number of preprints tends to increase with institution size for ECRs, although the interaction between ECR status and institution size was not significant. The United States and United Kingdom contributed the greatest number of preprints by early career researchers, whereas non-western countries contributed relatively fewer preprints. This research provides empirical evidence regarding motivations of preprint usage and barriers surrounding large-scale adoption of preprinting in ecology and evolution.

Introduction

Preprints are free, publicly accessible, early versions of scientific articles. They are posted online prior to or in parallel with the peer review process and help shorten the temporal gap between completed studies and accessible research (Sarabipour et al., 2019; Vale, 2015). bioRxiv is one of the most popular preprint servers in the fields of ecology and evolution and hosts preprint articles with barrier-free access to manuscripts (Hyland, 2016; Merga & Mason, 2020). Preprints usually appear on bioRxiv within 48 hours (https://www.biorxiv.org/about/FAQ), whereas manuscripts accepted at the first journal they are submitted to take approximately four months to become visible, including the publication process (Himmelstein, 2015; Royle, 2015). It is not uncommon for manuscripts to be submitted consecutively to more than one journal and, as a result, the peer review process can take years (Cobb, 2017). The usage of preprints is likely driven by open access research availability and recognition: it demonstrates a paper is complete and ready for peer review. Preprints facilitate the sharing of knowledge prior to peer review and improve transparency through open access research. Increased availability and recognition of preprints has also led to increased citations (Serghiou & Ioannidis, 2018; Shuai et al., 2012), which is a key metric by which researchers are evaluated (Nicholas et al., 2019). Ultimately, preprints can be beneficial to all researchers, but the use of preprints might be especially beneficial to the unique challenges that Early Career Researchers (ECRs) face relative to senior researchers.

In this context, ECRs are defined as individuals who are at the beginning stages of their research careers and do not yet have established research programs or tenured positions (Laudel & Gläser, 2008; Nicholas et al., 2019). This cohort includes graduate students undertaking masters or doctoral degrees, postdoctoral researchers, and un-tenured professors. In contrast, senior researchers are individuals who have held independent academic positions for greater than five years. While both groups hold research positions and are members of the scientific community, they face different challenges when publishing peer reviewed research (Laudel & Gläser, 2008). Peer review is the process of subjecting an author's scholarly work, research, or ideas, to the scrutiny of others who are experts in the field (Kelly et al., 2014). It is the standardized procedure by which the validity of novel scientific research is probed prior to publication and dissemination. The peer review process requires that several relevant but impartial experts in the field closely examine the manuscript and determine its value to the scientific community. The value of peer review lies in its responsibility to determine the importance and originality of research, as well as identify any scientific or methodological errors. As such, the peer review process can be lengthy and even overwhelming for researchers, especially for ECRs who face compounding challenges due to their career stage.

The usage of preprints could provide some relief to these challenges. Preprints can help counteract the collaborative, financial, and time constraints that ECRs face and provide an opportunity for them to gain feedback from peers in a more cost-effective manner (Merga & Mason, 2020). ECRs do not typically have the same funding levels as established researchers and may not be able to share their research in open access journals that allow for increased visibility and discussion of their research (Merga & Mason, 2020). Publishing articles in open access often costs thousands of dollars, which is prohibitive for many ECRs lacking financial support of their own or their mentor's, whereas the per-paper processing costs of preprints are low and the cost to researchers is absent (Sarabipour et al., 2019). This lower cost makes preprints potentially vital to ECRs as they provide opportunities to increase the volume and quality of free and informal feedback and collaboration on a greater scale relative to the typical peer review process that averages just over two formal reviews (Huisman & Smits, 2017; Penfold & Polka, 2020; Sarabipour et al., 2019). Additionally, the costs associated with open access publishing may be the reason that small institutions publish proportionately fewer open access articles (Shafer, 2020). These disparities emphasize the discriminatory nature of open access processing charges against authors with little access to funding (Burchardt, 2014) and the disproportionate benefit to preprint usage among institutions of different sizes. Ultimately, this literature highlights the financial burden of publishing research in open access journals and provides further motivation for the use of a low-cost preprint journal.

There is some resistance to preprinting, especially as perceived by senior researchers, as no formal peer review takes place and the onus is on the reader to interpret the accuracy and significance of the findings (Bove-Fenderson et al., 2018; Fry et al., 2019). However, making research readily available while it undergoes peer review facilitates early public access to novel data and methodologies that can inform ECRs' decisions regarding their own research, saving time and money. Employment in the fields of ecology and evolution requires quality scholarly research outputs for career advancement, which can be a challenge for ECRs (Hyland, 2016;

Merga & Mason, 2020). While an impressive publication record is important at all career stages, the increased stakes due to lower levels of job security for ECRs and the typically limited time window for applying for scholarships, grants, or faculty positions make this delay between submission and publication particularly detrimental for individuals at this stage. Preprints can reduce these barriers by allowing ECRs to make their work publicly available more rapidly and at no cost, thus increasing research visibility (Serghiou & Ioannidis, 2018) and ultimately assisting career development (Berg et al., 2016).

As masters and PhD programs are often short term (typically 2-5 years; (DeClou, 2017)), there is an especially large pressure on graduate students to conduct and publish high quality research in a short period of time (Browning et al., 2017). ECRs also face challenges surrounding financial and employment instability (DeClou, 2017; Nicholas et al., 2017). These challenges are faced less often by senior researchers, who typically have tenure or seniority at their organization (McAlpine & Amundsen, 2015). The challenges of writing, managing journal requirements, and dealing with the academic review process are shared among all researchers. However, the experience is likely different between ECRs and non-ECRs, as the familiarity with peer review and publication differs (Nicholas et al., 2017). Academic positions are becoming more competitive over time, and because ECRs are still developing their research niche, they often face a limited amount of support in the form of experienced colleagues and funding, both of which senior researchers tend to gain over time (Bazeley, 2003; McAlpine & Amundsen, 2015). Concurrently, there is increasing demand to demonstrate that research results in social, economic, and policy change (Chikoore et al., 2016). These hurdles disproportionately affect ECRs because translating research into these arenas takes additional time, which is limited. However, preprints can potentially help to address some of these issues.

Herein, we examine journal articles in four popular journals in the fields of ecology and evolution and one preprint server to assess whether the usage of preprints differs based on career stage. Due to the challenges that ECRs face and the benefits that preprints might provide, we hypothesize that ECRs disproportionately utilize preprint servers relative to senior researchers. We also considered the effect of institution size on preprint rates, because of its positive association with funding availability and open access publication rates (Shafer, 2020), and considered potential disparities between countries. We predicted that, when either the first or last author was identified as an ECR, it would have a positive impact on the number of preprints by the first author, as ECRs would be more likely to use preprint servers. The discussion regarding the benefit of increased preprint prevalence is ongoing (see Harmit Malik Tweet, 2020). Among others, the journal eLife proposes a shift from the typical peer review process to a more open and transparent framework, where journals help transform preprints into high-quality published manuscripts (eLife Tweet, 2021; Michael Eisen Tweet, 2020). We shed light on some of the factors involved in such a transition.

Methods

Data collection

Articles from two open access journals (Ecology and Evolution and PLOS One), two subscription-based journals (Proceedings of the Royal Society B and Ecology), and one preprint server (bioRxiv) were collected from 2019, using the search term 'ecology and evolution' in each journal's webpage. The first 100 full-text English articles from each journal were collected to produce a final dataset of 200 open access articles, 200 closed access articles, and 100 preprints. Fewer than 100 articles from Proceedings of the Royal Society B in 2019 fit our criteria, so the search was expanded to include publications in 2020 to achieve the desired sample size. The title of the article, number of authors, names of first and last authors, and author affiliations were collected. The H index and total number of publications (considered here as all entries on a google scholar profile) were determined for the first and last authors using Google Scholar. Articles were excluded if the first or last author did not have a Google Scholar profile. The bioRxiv database was used to determine the number of preprints an author has submitted. We defined ECR as a student at any stage (Laudel & Gläser, 2008), as well as individuals that held an independent academic position for five years or less (NSERC, 2020). The ECR status of the first and last authors were determined by examining institutional profiles and websites, Google Scholar, or personal websites, in that order. Articles were excluded if the ECR status could not be determined. Lastly, information on the affiliations of the first and last authors, including size, type, and country were collected. Institutions were quantified as small (<10,000 students), medium (10,001-19,999 students), or large (>20,000 students), following Shafer (2020).

Data analysis

The base package in R v 4.0.2 was used to generate models and statistical analysis. A generalized linear model (GLM) was run, with the number of preprint articles as a response variable. Fixed explanatory variables included the career status of both the first and last author (factor), the institution size of the first author (factor), and an interaction term of the career status and institution size of the first author. The total number of publications by the first author was log transformed and included as an offset variable, so that we were able to effectively model the relationship as a rate and use the Poisson distribution (Shafer, 2020). A Poisson GLM indicated that there was overdispersion in the data. We corrected the standard errors using a quasi-GLM model where the variance is given by $\Psi \times \mu$, μ is the mean, and Ψ is the dispersion parameter (Zuur et al., 2019).

To determine if an individual who publishes more generates more preprints, we tested for a correlation between the number of preprints and total publications of first and last authors using a Spearman's rank correlation test. Furthermore, we completed a one-way analysis of variance (ANOVA) to determine if the mean number of preprints differed by the country associated with the first author. We then used ArcGIS Pro (version 2.6.3) to visualize the distribution of preprints and proportion of early career researchers by country. The percent of total preprints represents the sum of first author preprints in each country divided by the total number of preprints counted in this study. The proportion ECR represents the proportion of preprints in each country with an ECR as the first author. A value of 1 indicates all preprints were submitted by an ECR and a value of 0 indicates that all preprints were submitted by a senior researcher.

Results

A total of 500 articles were included in our analysis (Proceedings of the Royal Society B = 100 [N_{2019} = 49, N_{2020} = 51], PLoS ONE = 100, Ecology = 100; Ecology and Evolution = 100, bioRxiv = 100). The corrected GLM demonstrated that career status of both the first and last author had significant effects on the number of preprints generated by the first author (Fig. 1a-b, Table 1-2). When either the first or last author was not an early-career researcher, there were relatively fewer preprints generated, while this effect was greater when the first author was not an ECR, indicated by the lower incidence rate ratio (Table 2). There is also a marginal effect of first author institution size on preprint usage whereby the number of preprints tends to increase with institution size (Table 2), although this is only evident for ECRs (Fig. 1c) and the interaction between ECR status and institution size was not significant (Table 1-2). Ultimately, this indicates that individuals in earlier stages of their careers generate more preprints relative to non-ECRs, with institution size also possibly playing a role in this relationship (Fig. 1).

Spearman's rank correlation tests indicated no correlation between the number of preprints and total number of publications for the first author ($r = 0.031$, $p = 0.50$) or last author ($r = 0.050$, $p = 0.28$). The United States and United Kingdom contributed the greatest number of preprints by early career researchers, whereas non-western countries contributed relatively fewer preprints (Fig. 2). However, there was no significant difference in mean number of preprints of the first author between countries ($F_{39,401} = 1.177$, $p = 0.22$).

Discussion

Preprints are used differentially by researchers at different career stages and varies with institution size. ECRs generate more preprints than non-ECRs and preprint usage tends to increase with the institution size of first authors. Preprints are already benefiting ECRs and the broader scientific community (Penfold & Polka, 2020), but the large-scale adoption of preprints across all career stages and institution sizes is necessary for these benefits to be distributed equitably. A strong preprint culture has the potential to reduce the negative impacts of the current publishing landscape with respect to the lengthy review process and financial burdens associated with open access research that are disproportionately shouldered by ECRs.

The usage of preprint servers has many advantages for ECRs, including the rapid dissemination of research, broad visibility of research output through open access, and more inclusive and transparent peer review, as feedback can be obtained from the entire scientific community rather than a few reviewers (Desjardins-Proulx et al., 2013; Penfold & Polka, 2020; Sarabipour et al., 2019). While the value of a preprint is difficult to compare to a peer reviewed manuscript, developing an environment where preprinting is the norm allows individuals to disseminate their research and build a reputation while reducing the likelihood that their work is not recognized by grant or job evaluation committees (Desjardins-Proulx et al., 2013).

Senior authors often play a crucial role with respect to an ECR's opportunity to publish in a journal with a moderate-to-high impact factor (Sekara et al., 2018), whereas preprints offer a space that is free of publication bias and encourages sharing of diverse researchers' works, at any career stage (Jennions & Møller, 2002; Sarabipour et al., 2019). However, ECRs typically have less experience, power, and security relative to senior researchers and, thus, exert less control over the decision to utilize a preprint server. As such, the balance of power often lies with the senior researcher, wherein they may not support an ECR to utilize a preprint server. It is possible that the senior researcher may perceive preprinting to offer no benefit or even cause harm to themselves, their trainee(s), or the field at large. For example, concerns about being "scooped" (i.e., competing researchers using knowledge gained from the preprint to publish similar findings in the peer-reviewed literature before the preprint authors) have been raised, though evidence is lacking to support this as a substantial risk (Penfold & Polka, 2020; Sever et al., 2019). Related to this are concerns that preprints, although assigned a Digital Object Identifier (DOI) establishing a permanent record of submission date and content, will not be sufficient for establishing precedence of discoveries (Sever et al., 2019). This has not been widely tested to our knowledge. Finally, there are concerns that preprint servers provide a means of widely disseminating poor quality research, which can then be cited, thus threatening the integrity of the scientific literature (Maggio et al., 2018). Senior researchers might choose to place these concerns above the potential benefits for ECRs. However, the interplay between ECRs and senior researchers in regard to the usage of preprint servers is likely complex. ECRs might also be skeptical to use preprint servers if they are insecure about their research being made public prior to being vetted by impartial experts. This concern might be especially prevalent when ECRs engage with a new subject, stray from the research area of their advisor(s) or belong to underrepresented or marginalized groups that disproportionately suffer from "imposter syndrome" in academia (Pulliam & Gonzalez, 2018; Bravata et al., 2020). While we show that the ECR status of first and last authors influences preprint usage, the decision-making process behind this trend should be explored in future research.

Smaller institutions publish fewer open access articles relative to researchers at large universities (Shafer, 2020), a trend that seems to hold true for preprints, despite their accessibility. Further, data from Robinson-Garcia et al. (2020) suggests that countries with higher rates of open access publications than the median also exhibit greater rates of preprint usage. This might be detrimental for ECRs at smaller institutions, as it could further limit their research impact and would place ECRs at small universities in countries with low open access rates at an even greater disadvantage from a research impact point of view. Further, our results uncovered geographical heterogeneity in the rate of preprint usage (Fig. 2). South American and Asian countries published fewer preprints compared to North American and European countries and Australia (Fig. 2). Some of the top countries in terms of open access publications rates, outlined by Robinson-Garcia et al. (2020), such as the United Kingdom, United States, France, and Spain, also produced some of the highest rates of preprint usage. For the most part, countries with lower open access publication rates also exhibited the lowest preprint usage rates. An exception to the above is Australia, which contributed a high number of preprints while posting an open access publication rate slightly below the global median (Robinson-Garcia et al., 2020). Our results demonstrated that preprint publications within a country mostly mimic that country's open access publication landscape.

Although not examined directly in the present study, the potential conflicts and disparities surrounding preprinting are likely further compounded among ECRs who belong to underrepresented groups. Underrepresented and historically marginalized groups face additional burdens in advancing in the fields of ecology and evolution (Fox et al., 2018, 2019; Fox & Paine, 2019; Miriti et al., 2020; Schell et al., 2020; Tseng et al., 2020). In particular, women and Black, Indigenous, and People of Color (BIPOC) are underrepresented in ecology and evolution, especially in positions of power (O'Brien et al., 2020; Wei et al., 2020). In the United States, between 2014 and 2018, only 1.2% of ecology and evolutionary biology PhD graduates identified as Black or Indigenous. This contrasts with the 16% of the American population identifying with those demographics (Tseng et al., 2020). The lack of ethnic diversity in ecology and evolution can be attributed to a variety of factors including systemic disparities in education between communities and a lack of role models leading to a reduced sense of belonging in the field for non-white students (O'Brien et al., 2020). Due to the discrepancies between the general population and demographics of the field, it would be remiss to assume that all ECRs face the same battles. The accumulation of challenges due to the intersectionality of career status and other social identities further amplifies the difficulties that many individuals experience on a daily basis (Crenshaw 1989, Wanelik et al. 2020). There are barriers at every career stage and the disparity between ECRs and non-ECRs in preprint usage potentially exacerbates the difficulties that underrepresented ECRs face, which is a topic that warrants further study.

In addition to the challenges faced by ethnic minorities, women are also underrepresented in this field. For example, in ecology and zoology, women represent less than one-third of authors and research groups led by women published with >60% female coauthors, whereas in male-led groups, <20% of coauthors were female (Salerno et al., 2019). We initially considered exploring the relationship between ECR status and preprint usage among individuals of different genders within this study. However, without data on self-reported gender identities, we were unable to do so in a way that does not risk mis-identifying and causing harm to individuals (e.g., by using gender assignment tools based on first names) (Cameron & Stinson, 2019). While gender is an extremely important factor to analyze in respect to research trends in ecology and evolution, until there is a method by which individuals can self-identify their preferred gender/pronouns, we do not feel it is appropriate to use tools that may misassign gender and cause harm. Subsequently, we argue that this provides motivation for authors to include their self-identified pronouns in research profiles, biographies, and publications (as we have done in the present work), so that relationships of these kind can be explored whilst ensuring as safe a space as possible.

Conclusion

This research demonstrates that preprint servers are used disproportionately between ECRs and non-ECRs and between institutions of different sizes. We suggest key factors that may lead to a differential usage of preprints among researchers of varying career stages and discuss the effects that using preprint servers can have on career development. The shift among researchers to use preprint servers may simply be one of progressiveness and the natural development of the field of ecology and evolution. However, we noted that there is a myriad of benefits to using preprint servers that are specifically valuable for ECRs, which may be the reason behind the marked shift in preprint usage. Open access research is associated with increases in citations, media attention, potential collaborators, job opportunities, and funding opportunities (Fu & Hughey, 2019; McKiernan et al., 2016). These benefits might drive ECRs to make their research publicly accessible and, in turn, be linked to the greater usage of preprint servers by ECRs.

This research provides evidence for the unequal usage of preprint servers among researchers of varying career stages and is necessary to facilitate further discussion surrounding the larger-scale adoption of preprints in the field of ecology and evolution. Preliminary discussions have been had regarding the adoption of mandatory usage of preprint servers prior to peer review (Desjardins-Proulx et al., 2013; Harmit Malik Tweet, 2020; Moi Exposito-Alonso Tweet, 2020; Vince Buffalo Tweet 2021). It is evident that career stage influences preprint usage, and due to the multitude of benefits, we believe further discussions and studies of this kind are necessary to address the unique needs of ECRs in the field of ecology and evolution with respect to preprint usage. This research aims to illuminate the landscape of preprint servers specifically in ecology and evolution, while future research should aim to determine the cause(s) of disproportionate usage of preprints by ECRs and larger institutions to help reduce possible barriers to preprinting for other groups. Ultimately, a strong and widely adopted preprint culture in ecology and evolution may help facilitate greater preprint usage among historically marginalized groups, aiding in career development for those who are underrepresented in the field.

Literature Cited

Bazeley P. 2003 Defining'early career'in research. *High. Educ*. 45, 257–79. https://doi.org/10.1023/A:1022698529612

Berg BJM, Bhalla N, Drubin G, Fraser JS, Carol W. 2016 Preprints for the life sciences. *Science* **352**, 899–901. https://doi.org/10.1126/science.aaf9133

Bove-Fenderson E, Duffy K, Mannstadt M. 2018 Broadening Our Horizons: JBMR and JBMR Plus Embrace Preprints. *J. Bone Miner. Res.* **33**, 185–7. https://doi.org/ 10.1002/jbm4.10042

Bravata D, Madhusudhan D, Boroff M, Cokley K. 2020 Commentary: Prevalence, Predictors, and Treatment of Imposter Syndrome: A Systematic Review. *J. Ment. Heal. Clin. Psychol.* **4**, 12–6.

Browning L, Thompson K, Dawson D. 2017 From early career researcher to research leader: survival of the fittest? *J. High. Educ. Policy Manag.* **39**, 361–77. https://doi.org/10.1080/1360080X.2017.1330814

Burchardt J. 2014 Researchers Outside APC-Financed Open Access. *SAGE Open* **4**, 1-11. https:// doi.org/10.1177/2158244014551714

Cameron JJ, Stinson DA. 2019 Gender (mis)measurement: Guidelines for respecting gender diversity in psychological research. *Soc. Personal. Psychol. Compass.* 13, 1–14. https://doi.org/10.1111/spc3.12506

Chikoore L, Probets S, Fry J, Creaser C. 2016 How are UK Academics Engaging the Public with their Research A Cross-Disciplinary Perspective. *High. Educ. Q.* **70**, 145–69. https://doi.org/10.1111/hequ.12088

Cobb M. 2017 The prehistory of biology preprints: A forgotten experiment from the 1960s. *PLoS Biol.* **15**, 1–12. https://doi.org/10.1371/journal.pbio.2003995

Crenshaw, K. 1989 Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. University of Chicago Legal Forum. 139–167.

DeClou L. 2017 Who Stays and for How Long: Examining Attrition in Canadian Graduate Programs. *Can. J. High. Educ.* **46**, 174–98.

Desjardins-Proulx P, White EP, Adamson JJ, Ram K, Poisot T, Gravel D. 2013 The Case for Open Preprints in Biology. *PLoS Biol.* **11**, 3–7. https://doi.org/10.1371/journal.pbio.1001563

eLife (2021, January 18). [Tweet]. https://twitter.com/elife/status/1351202920205869057?s=12

Fox CW, Duffy MA, Fairbairn DJ, Meyer JA. 2019 Gender diversity of editorial boards and gender differences in the peer review process at six journals of ecology and evolution. *Ecol. Evol*. **9**, 13636–49. https://doi.org/10.1002/ece3.5794

Fox CW, Paine CET. 2019 Gender differences in peer review outcomes and manuscript impact at six journals of ecology and evolution. *Ecol. Evol*. **9**, 3599–619. https://doi.org/10.1002/ece3.4993

Fox CW, Ritchey JP, Paine CET. 2018 Patterns of authorship in ecology and evolution: First, last, and corresponding authorship vary with gender and geography. *Ecol. Evol*. **8**, 11492–507. https://doi.org/10.1002/ece3.4584

Fry NK, Marshal H, Mellins-Cohen T. 2019 In praise of preprints. *Microb. Genomics* 5, 5–7. https://doi.org/10.1099/mgen.0.000259

Fu DY, Hughey JJ. 2019 Releasing a preprint is associated with more attention and citations for the peer-reviewed article. *Elife* **8**, 1–12. https://doi.org/ 10.7554/eLife.52646

HarmitMalik (2020, December 21). [Tweet]. https://twitter.com/harmitmalik/status/1340932602459353088?s=12

Himmelstein D. Publication delays at PLOS and 3,475 other journals. Satoshi Village [Internet]. 2015. Available from: https://web.archive.org/web/20180503071358/http://blog.dhimmel.com/plos-and-publishing-delays/. [cited 2021 Februrary 21].

Huisman J, Smits J. 2017 Duration and quality of the peer review process: the author's perspective. *Scientometrics* **113**, 633–50. https://doi.org/10.1007/s11192-017-2310-5

Hyland K. 2016 Academic publishing and the myth of linguistic injustice. *J. Second Lang. Writ.* **31**, 58–69. https://doi.org/10.1016/j.jslw.2016.01.005

Jennions MD, Møller AP. 2002 Publication bias in ecology and evolution: An empirical assessment using the "trim and fill" method. *Biol. Rev. Camb. Philos. Soc.* **77**, 211–22. https://doi.org/ 10.1017/s1464793101005875

Kelly J, Sadeghiegh T, Adeli K. 2014 Peer review in scientific publications: benefits, critiques, & a survival guide. *eJIFCC* **25**, 227–43.

Laudel G, Gläser J. 2008 From apprentice to colleague: The metamorphosis of Early Career Researchers. *High. Educ.* **55**, 387–406. https://doi.org/10.1007/s10734-007-9063-7

Maggio LA, Artino AR, Driessen EW. 2018 Preprints: Facilitating early discovery, access, and feedback. *Perspect. Med. Educ.* **7**, 287–9. https://doi.org/ 10.1007/s40037-018-0451-8

McAlpine L, Amundsen C. 2015 Early career researcher challenges: substantive and methods-based insights. *Stud. Contin. Educ.* **37**, 1–17. https://doi.org/10.1080/0158037X.2014.967344

McKiernan EC, Bourne PE, Brown CT, Buck S, Kenall A, Lin J, et al. 2016 How open science helps researchers succeed. *Elife* **5**, 1–19. https://doi.org/10.7554/eLife.16800.001

Mbeisen. (2020, December 1). [Tweet]. https://twitter.com/mbeisen/status/1333852296581550080

Merga M, Mason S. 2020 Sharing research with academia and beyond: Insights from early career researchers in Australia and Japan. *Learn. Publ.* **33**, 277–86. https://doi.org/10.1002/leap.1296

MExpositoAlonso (2020, May 28). [Tweet]. https://twitter.com/MExpositoAlonso/status/1262585395662819328

Miriti MN, Bailey K, Halsey SJ, Harris NC. 2020 Hidden figures in ecology and evolution. *Nat. Ecol. Evol.* **4**, 1282. https://doi.org/10.1038/s41559-020-1270-y

Nicholas D, Rodríguez-Bravo B, Watkinson A, Boukacem-Zeghmouri C, Herman E, Xu J, et al. 2017 Early career researchers and their publishing and authorship practices. *Learn. Publ.* **30**, 205–17. https://doi.org/10.1002/leap.1102

Nicholas D, Watkinson A, Boukacem-Zeghmouri C, Rodríguez-Bravo B, Xu J, Abrizah A, et al. 2019 So, are early career researchers the harbingers of change? *Learn. Publ.* **32**, 237–47. ttps://doi.org/10.1002/leap.1232

NSERC. 2020 *NSERC - what is an early career researcher?* Available from: https://www.nserc-crsng.gc.ca/Professors-Professeurs/Grants-Subs/DGCategories-SDCategories_eng.asp [cited 2020 September 4].

O'Brien LT, Bart HL, Garcia DM. 2020 Why are there so few ethnic minorities in ecology and evolutionary biology? Challenges to inclusion and the role of sense of belonging. *Soc. Psychol. Educ.* **23**, 449–77. https://doi.org/10.1007/s11218-019-09538-x

Penfold NC, Polka JK. 2020 Technical and social issues influencing the adoption of preprints in the life sciences. *PLoS Genet.* **16**, 1–16. https://doi.org/10.1371/journal.pgen.1008565

Pulliam N, Gonzalez CE. 2018 Success or Fraud? Exploring the Impacts of the Impostor Phenomenon Among High Achieving Racial/Ethnic Minority and First-Generation College Students. *J. Access, Retention, Incl. High. Educ.* **1**, 33–45.

Robinson-Garcia N, Costas R, van Leeuwen TN. 2020 Open Access uptake by universities worldwide. *PeerJ* **2020**, 1–20. https://doi.org/ 10.7717/peerj.9410

Royle S. Waiting to happen II: Publication lag times [Internet]. Available from:http://web.archive.org/web/20190412012435/http://quantixed.org/2015/03/16/waiting-to-happen-ii- publication-lag-times/. [cited 2021 February 21].

Salerno PE, Páez-Vacas M, Guayasamin JM, Stynoski JL. 2019 Male principal investigators (almost) don't publish with women in ecology and zoology. *PLoS One* **14**, 1–14. https://doi.org/10.1371/journal.pone.0218598

Sarabipour S, Debat HJ, Emmott E, Burgess SJ, Schwessinger B, Hensel Z. 2019 On the value of preprints: An early career researcher perspective. *PLoS Biol.* **17**, 1–12. https://doi.org/10.1371/journal.pbio.3000151

Schell CJ, Guy C, Shelton DS, Campbell-Staton SC, Sealey BA, Lee DN, et al. 2020 Recreating Wakanda by promoting Black excellence in ecology and evolution. *Nat. Ecol. Evol*. https://doi.org/10.1038/s41559-020-1266-7

Sekara V, Deville P, Ahnert SE, Barabási AL, Sinatra R, Lehmann S. 2018 The chaperone effect in scientific publishing. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 12603–7. https://doi.org/10.1073/pnas.1800471115

Serghiou S, Ioannidis JPA. 2018 Altmetric Scores, Citations, and Publication of Studies Posted as Preprints. *JAMA - J. Am. Med. Assoc.* **319**, 402–3. https://doi.org/ 10.1001/jama.2017.21168

Sever R, Roeder T, Hindle S, Sussman L, Black KJ, Argentine J, et al. 2019 BioRxiv: The preprint server for biology. *bioRxiv* 1–19. https://doi.org/10.1101/833400

Shafer ABA. 2020 Are we paying-to-play? A quantitative assessment of Canadian Open Access research in ecology and evolution. *Facets*. https://doi.org/ 10.1139/facets-2020-0040

Shuai X, Pepe A, Bollen J. 2012 How the Scientific Community Reacts to Newly Submitted Preprints: Article Downloads, Twitter Mentions, and Citations. *PLoS One* **7**, 1–8. https://doi.org/10.1371/journal.pone.0047523

Tseng M, El-Sabaawi RW, Kantar MB, Pantel JH, Srivastava DS, Ware JL. 2020 Strategies and support for Black, Indigenous, and people of colour in ecology and evolutionary biology. *Nat. Ecol. Evol.* **4**, 1288–90. https://doi.org/10.1038/s41559-020-1252-0

Vale RD. 2015 Accelerating scientific publication in biology. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 13439–46. https://doi.org/10.1073/pnas.1511912112

Vsbuffalo (2021, February 9). [Tweet]. https://twitter.com/vsbuffalo/status/1359301637035094016?s=12

Wanelik KM, Griffin JS, Head ML, Ingleby FC, Lewis Z. 2020 Breaking barriers? Ethnicity and socioeconomic background impact on early career progression in the fields of ecology and evolution. *Ecol. Evol.* **10**, 6870–80. https://doi.org/10.1002/ece3.6423

Wei Q, Lachapelle F, Fuller S, Corrigall-Brown C, Srivastava DS. 2020 Working groups, gender and publication impact of Canada's ecology and evolution faculty. *bioRxiv*. https://doi.org/10.1101/2020.05.12.092247

Zuur, Alain F, Ieno, Elena N, Walker, Neil J, Saveliev, Anatoly A, Smith, Graham M. 2019.

GLM and GAM for Count Data. Mixed effects models and extensions in ecology with R p. 209–39 In New York, NY: Springer Science.

Table 1. Deviance table from generalized linear models for the effects of first author Early Career Researcher (ECR) status, last author ECR status, and first author institution size on the number of preprints generated by the first author. The P-values were obtained from $\chi 2$ tests of whether the model fit improved by sequentially adding first author ECR status, last author ECR status, first author institution size, and the interaction between first author ECR status of the first author and first author institution size to the null model. Significance values are in bold and were determined using $\alpha = 0.05$

Model term	d.f.	Deviance	Residual d.f.	Residual deviance	*P*-value
null			498	2102.5	
First author ECR status	1	416.94	497	1685.5	**$< 2.2 \times 10^{-16}$**
Last author ECR status	1	28.319	496	1657.2	**0.002**
First author institution size	2	29.538	494	1627.7	0.057
First author ECR status × First author institution size	2	2.5912	492	1625.1	0.78

Table 2. Results from a quasipoisson family generalized linear model, where the number of preprint articles of the first author is the response variable. Significance values are in bold and were measured using α = 0.05. The intercept in this model represents early career researchers at large institutions.

	Incidence Rate Ratios	Confidence interval	*P*-value
(Intercept)	0.12	0.08 – 0.17	**<0.001**
First author status: non-early career researcher	0.20	0.12 – 0.30	**<0.001**
Last author status: non-early career researcher	0.64	0.44 – 0.95	**0.024**
Medium-sized institution	0.77	0.43 – 1.30	0.361
Small-sized institution	0.55	0.28 – 0.97	0.056
First author status: non-early career researcher * Medium-sized institution	1.42	0.52 – 3.65	0.475
First author status: non-early career researcher * small-sized institution	1.08	0.32 – 3.24	0.892

Figure Legends

Figure 1. The raw number of preprint articles produced by (A) first and (B) last authors of ecology and evolution articles (n=500) according to Early Career Researcher (ECR) status. (C) Model-corrected values for the number of preprint articles produced according to ECR status and institution size of the first author. Lines in the violins represent the 25%, 50%, 75%, and 95% quartiles.

Figure 2. Map showing the percent of preprints (n=734) and proportion of Early Career Researcher (ECR) authors by country. Point size indicates the percentage of the total preprints by first authors assigned to each country. Color indicates the proportion of preprints in each country that were classified as ECRs.

Data Accessibility Statement

The raw data and R scripts used to perform analyses and generate figures for this manuscript are available at https://github.com/geneticsjesse/ECR_manuscript

Competing Interests Statement

The authors declare that they have no competing interests.

Author Contributions

J.F.W. conceived the ideas; J.F.W., L.M, S.E.H., M.L.C., M.N.D., K.B.Y., and C.I.E. collected materials and data; J.F.W. analysed the data with input from R.A.O.; J.F.W. wrote the paper with contributions from all co-authors; and R.A.O. provided edits, feedback, and advisory assistance.

Acknowledgements

Special thanks to Kathleen Lo and Dr. Aaron B.A. Shafer for their comments on earlier drafts of this manuscript. Thanks to the Wildlife and Applied Genomics Lab at Trent University and specifically Dr. A.B.A Shafer for discussions surrounding methodologies and our experimental approach with specific respect to their paper *Shafer, A. B. A. (2020). Are we paying-to-play? A quantitative assessment of Canadian Open Access research in ecology and evolution.* S.E. Haworth was supported by an NSERC-CGSM, C.I. Elliott was supported by an NSERC-USRA, and M.L. Cossette was supported by an NSERC-USRA and an OGS scholarship. R.A. Oomen was supported by the James S. McDonnell Foundation.