AUTHOR	Daniel W. A. Noble, Zoe A. Xirocostas, Nicholas C. Wu, April Robin Martinig, Rafaela A. Almeida, Kevin R. Bairos-Novak, Heikel Balti, Michael G. Bertram, Louis Bliard, Jack A. Brand, Ilha Byrne, Ying-Chi Chan, Dena Jane Clink, Quentin Corbel, Ricardo A. Correia, Jordann Crawford-Ash, Antica Culina, Elvira D'Bastiani, Gideon G. Deme, Melina deSouza Leite, Félicie Dhellemmes, Shreya Dimri, Szymek M. Drobniak, Alexander D. Elsy, Susan E. Everingham, Samuel J. L. Gascoigne, Matthew J. Grainger, Gavin C. Hossack, Knut Anders Hovstad, Edward R. Ivimey-Cook, Matt Lloyd Jones, Ineta Kačergytė, Georg Küstner, Dalton C. Leibold, Magdalena M. Mair, Jake Martin, Ayumi Mizuno, Iain R.Moodie, David Moreau, Rose E. O'Dea, James A. Orr, Matthieu Paquet, Rabindra Parajuli, Joel L. Pick, Patrice Pottier, Marija Purgar, Pablo Recio, Dominique G. Roche, Raphaël Royauté, Saeed Shafiei Sabet, Julio M. G. Segovia, Inês Silva, Alfredo Sánchez-Tójar, Bruno E. Soares, Birgit Szabo, Elina Takola, Eli S. J. Thoré, Bishnu Timilsina, Natalie E. van Dis, Wilco C. E. P. Verberk, Stefan J. G. Vriend, Kristoffer H. Wild, Coralie Williams, Yefeng Yang, Shinichi Nakagawa and Malgorzata Lagisz
TITLE	The promise of community-driven preprints in ecology and evolution
YEAR	2025
DOI	https://doi.org/10.1098/rspb.2024.1487
VERSION	Author's accepted manuscript
CITATION	Noble DWA et al. 2025, The promise of community-driven preprints in ecology and evolution. Proc. R. Soc. B 292: 20241487. https://doi.org/10.1098/rspb.2024.1487

The promise of community-driven preprints in ecology and evolution

Authors: Daniel W.A. Noble[1], Zoe A. Xirocostas[2], Nicholas C. Wu[3], April R. Martinig[4], Rafaela A. Almeida[5], Kevin R. Bairos-Novak[6], Heikel Balti[7], Michael G. Bertram[8,9,10], Louis Bliard[11], Jack A. Brand[8], Ilha Byrne[13], Ying-Chi Chan[14], Dena J. Clink[15], Quentin Corbel[16], Ricardo A. Correia[17], Jordann Crawford-Ash[18], Antica Culina[19], Elvira D'Bastiani[20], Gideon G. Deme[21,22], Melina de Souza Leite[23], Félicie Dhellemmes[24,25], Shreya Dimri[26], Szymek M. Drobniak[4,27], Alexander D. Elsy[28], Susan E. Everingham[29], Samuel J. L. Gascoigne[30], Matthew J. Grainger[31], Gavin C. Hossack[32], Knut Anders Hovstad[33], Ed R. Ivimey-Cook[34], Matt Lloyd Jones[35], Ineta Kačergytė[36], Georg Küstner[37], Dalton C. Leibold[1], Magdalena M. Mair[38], Jake Martin[8,9,39], Ayumi Mizuno[4, 40], Iain R. Moodie[41], David Moreau[42], Rose E. O'Dea[43], James A. Orr[44], Matthieu Paquet[45], Rabindra Parajuli[46], Joel L. Pick[47], Patrice Pottier[1,4], Marija Purgar[19], Pablo Recio[1], Dominique G. Roche[48], Raphaël Royauté[49], Saeed Shafiei Sabet[50], Julio M. G. Segovia[26], Inês Silva[51], Alfredo Sánchez-Tójar[26], Bruno E. Soares[52], Birgit Szabo[53], Elina Takola[54], Eli S.J. Thoré[8,55], Bishnu Timilsina[56], Natalie E. van Dis[57], Wilco C.E.P. Verberk[58], Stefan J.G. Vriend[59], Kristoffer H. Wild[1,60], Coralie Williams[4], Yefeng Yang[4], Shinichi Nakagawa[4, 61]*, Malgorzata Lagisz[4,]*

Affiliations

[1]Division of Ecology and Evolution, Research School of Biology, The Australian National University, Canberra, ACT 2600, Australia
[2] School of Life Sciences, Faculty of Science, University of Technology Sydney, NSW 2007, Australia
[3]Hawkesbury Institute for the Environment, Western Sydney University, Richmond, NSW 2753, Australia
[4]Evolution and Ecology Research Centre, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW 2052, Australia
[5]Laboratory of Freshwater Ecology, Evolution and Conservation, KU Leuven, Belgium
[6]School of the Environment, The University of Queensland, Brisbane 4072, Australia
[7] Chrono-environnement UMR 6249, CNRS, Université Bourgogne Franche-Comté, F-25000, Besançon, France
[8]Department of Wildlife, Fish, and Environmental Studies, Swedish University of Agricultural Sciences, Umeå 907 36, Sweden
[9]Department of Zoology, Stockholm University, Stockholm 114 18, Sweden
[10]School of Biological Sciences, Monash University, Melbourne 3800, Australia
[11] Department of Evolutionary Biology and Environmental Studies, University of Zurich, Switzerland
[12] Department of Wildlife, Fish, and Environmental Studies, Swedish University of Agricultural Sciences, Umeå, Sweden
[13] School of the Environment, The University of Queensland, Brisbane 4072, Australia
[14] Swiss Ornithological Institute, Sempach, Switzerland
[15] K. Lisa Yang Center for Conservation Bioacoustics, Cornell Lab of Ornithology, Cornell University, Ithaca, NY, USA
[16] Theoretical and Experimental Ecology Station (SETE), CNRS, France
[17] Biodiversity Unit, University of Turku, 20014 Turku, Finland
[18] Fenner School of Environment and Society, Australian National University, Canberra ACT
[19] Ruđer Bošković Institute, Zagreb, Croatia
[20] Department of Ecology and Evolutionary Biology, University of California, Los Angeles
[21] Department of Biology, Case Western Reserve University, Cleveland, OH 44106, USA
[22] Department of Science Laboratory Technology, University of Jos, Jos, Nigeria
[23] Department of Ecology, University of São Paulo, Brazil

[24] Cluster of Excellence "Science of Intelligence", Technical University of Berlin, Berlin, Germany
[25] Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany
[26] Department of Evolutionary Biology, Bielefeld University, Germany
[27] Institute of Environmental Sciences, Jagiellonian University, Krakow, Poland
[28] Department of Environmental Systems Science, ETH Zurich, Switzerland
[29] Institute of Plant Sciences and Oeschger Centre for Climate Change Research, University of Bern
[30] Department of Biology, University of Oxford, United Kingdom
[31] Norwegian Institute for Nature Research, Trondheim, Norway
[32] Port Perry, Ontario, Canada
[33] Norwegian Biodiversity Information Centre, Trondheim, Norway
[34] School of Biodiversity, One Health and Veterinary Medicine, University of Glasgow, UK
[35] European Centre for the Environment and Human Health, University of Exeter Medical School, Penryn, United Kingdom
[36] Ecology department, Swedish University of Agricultural sciences, Sweden
[37] Department of Animal Ecology and Tropical Biology, Biocenter, University of Würzburg, Würzburg, Germany
[38] Statistical Ecotoxicology, Bayreuth Center of Ecology and Environmental Research (BayCEER), University of Bayreuth, Germany
[39] School of Life and Environmental Sciences, Deakin University, Waurn Ponds, 3216, Australia
[40] Department of Biology, Faculty of Science, Hokkaido University, Japan
[41] Department of Biology, Lund University, Sweden
[42] Centre for Brain Research, University of Auckland, New Zealand
[43] School of Agriculture, Food and Ecosystem Sciences, University of Melbourne
[44] Department of Biology, University of Oxford
[45] Theoretical and Experimental Ecology Station (SETE), CNRS, France
[46] Department of Geosciences, Florida Atlantic University, Boca Raton, FL, USA
[47] Institute of Ecology and Evolution, University of Edinburgh, UK
[48] Department of Biology, Carleton University, Canada
[49] Université Paris-Saclay, INRAE, AgroParisTech, UMR EcoSys, Palaiseau, France
[50] Fisheries Department, Faculty of Natural Resources, University of Guilan, Swomeh Sara, Iran
[51] Center for Advanced Systems Understanding (CASUS), Helmholtz-Zentrum Dresden-Rossendorf (HZDR), Görlitz, Germany
[52] Institute of Environmental Change & Society, University of Regina
[53] Division of Behavioural Ecology, University of Bern, Switzerland
[54] Department of Computational Landscape Ecology, Helmholtz Center for Environmental Research - UFZ, Leipzig, Germany
[55] Laboratory of Adaptive Biodynamics, Research Unit of Environmental and Evolutionary Biology, Institute of Life, Earth, and Environment, University of Namur, Namur, Belgium
[56] The Arctic University Museum of Norway, The Arctic University of Norway (UiT), Norway
[57] Helsinki Institute of Life Sciences, Helsinki University, Finland
[58] Department of Ecology, Radboud University Nijmegen, Nijmegen
[59] Netherlands Institute of Ecology (NIOO-KNAW), Wageningen, Netherlands
[60] School of BioSciences, The University of Melbourne, Victoria 3010, Australia.
[61] Department of Biological Sciences, University of Alberta, CW 405, Biological Sciences Building, Edmonton, AB T6G 2E9, Canada

corresponding author, daniel.noble@anu.edu.au
* equal contribution
[5-60], Authors listed in alphabetical order

Abstract

Publishing preprints is quickly becoming commonplace in ecology and evolutionary biology. Preprints can facilitate the rapid sharing of scientific knowledge establishing precedence and enabling feedback from the research community before peer review. Yet, significant barriers to preprint use exist including language barriers, a lack of understanding about the benefits of preprints and a lack of diversity in the types of research outputs accepted (e.g., reports). Community driven preprint initiatives can allow a research community to come together to break down these barriers to improve equity and coverage of global knowledge. Here, we explore the first preprints uploaded to *EcoEvoRxiv* (n = 1216), a community-driven preprint server for ecologists and evolutionary biologists, to characterise preprint practices in ecology, evolution and conservation. Our perspective piece highlights some of the unique initiatives that *EcoEvoRxiv* has taken to break down barriers to scientific publishing by exploring the composition of articles, how gender and career stage influence preprint use, whether preprints are associated with greater open science practices (e.g., code and data sharing), and tracking preprint publication outcomes. Our analysis identifies areas that we still need to improve upon but highlight how community-driven initiatives, such as *EcoEvoRxiv*, can play a crucial role in shaping publishing practices in biology.

1. Introduction

Publishing preprints – papers communicating non-peer-reviewed research findings – is now an entrenched practice across a multitude of scientific disciplines [1]. Preprints in biology have had a slower uptake relative to other disciplines [2], but new discipline-specific preprint servers, such as *EcoEvoRxiv* (https://ecoevorxiv.org), provide a means by which ecologists and evolutionary biologists can disseminate research findings. Preprints attempt to break down barriers to scientific publishing by: 1) increasing the visibility of research and the speed at which research findings become available, which can lead to more citations [e.g., 3,4]; 2) helping establish the precedence of research findings; 3) removing financial barriers to open access publication; and 4) enabling earlier feedback from the research community [5–7]. Ultimately, preprints can facilitate the rapid sharing of scientific knowledge that can have significant impacts on fundamental and applied knowledge globally [8].

Preprint servers can empower researchers to make their research findings more accessible, open, and transparent but only if they are used as forums for spreading and discussing findings within a research community. However, significant barriers to the widespread adoption of preprints remain, ranging from a lack of clarity around preprint policies in journals [9] to a stigma within the research community that preprints are of poor quality [10] (but see [11]). Nonetheless, we lack an understanding of the factors that influence preprint use in ecology and evolution. Such an understanding may help improve current initiatives (see below), inform future ones and allow us to work harder in further breaking down barriers to scientific publishing.

EcoEvoRxiv is one of the few community-driven preprint servers that has paved the way for new initiatives, by accepting multilingual preprints, registered reports, and non-traditional research reports. Such initiatives are distinct from other preprint servers, such as *bioRxiv*, which only accept empirical research in English. In addition, community-driven servers like *EcoEvoRxiv* aggregate papers presenting research on similar topics, improving discoverability and opportunities for within-community debate compared with broader preprint servers. *EcoEvoRxiv* also promotes peer review and community discussion in the hopes of improving the quality of preprints and speeding up their peer-reviewed publication. For example, we encourage authors to use peer community review services such as Peer Community In (PCI) [12], which allow for fast, constructive peer review around a preprint with peer reviews being transparent and published online [12]. *EcoEvoRxiv* allows authors to submit both preprints and postprints (also known as author-accepted manuscripts). While preprints are versions of manuscripts posted by authors before peer-review, postprints are versions of peer-reviewed and accepted articles but without typesetting and formatting by a journal. The main reason for publishing postprints on a preprint server is to ensure published articles are openly accessible to everyone without a paywall (i.e., green open access). Even for articles published open access, depositing the postprint in a repository (e.g. Zenodo) or a preprint server strengthens permanence and access to the content of the article in the event of a journal's collapse or disappearance. Postprints can be published anytime, if journals allow it (which many do; see https://www.sherpa.ac.uk/romeo/).

Here, we explore the first preprints/postprints (n = 1216) uploaded to *EcoEvoRxiv* to characterise preprint practices in ecology and evolution. We aim to understand: 1) in what countries authors who use *EcoEvoRxiv* are located; 2) the taxonomic diversity of study systems used across articles; 3) whether preprint server use depends on career stage and gender; 4) the extent to which authors make use of preprint servers for reports and community-driven peer review; 5) the extent to which data and code are shared in preprints; and 6) how many preprints remain unpublished, and for those that are published, how long it took for them to become published. In the process, we also provide a summary of what makes *EcoEvoRxiv* distinct from other preprint servers to help further clarify the benefits of using community-driven preprint servers to disseminate research findings.

2. Getting to know your *EcoEvoRxiv* preprint server

EcoEvoRxiv is run by the Society for Open, Reliable, and Transparent Ecology and Evolutionary Biology (SORTEE)[13]. Originally launched in 2018 on the Center for Open Science preprint platform, *EcoEvoRxiv* has become a popular preprint server for ecologists and evolutionary biologists. The server has since been adopted by the California Digital Library (CDL). Editors are ecologists and evolutionary biologists from across the globe who volunteer their time to screen submitted papers and push new initiatives in the preprint space. To better understand preprint (and postprint) use on *EcoEvoRxiv*, we downloaded metadata on the accepted articles available on *EcoEvoRxiv* as of 2023-09-30 (see Supplement for more details on methods). We consider both preprints and postprints as 'articles'. After removing five duplicate titles (suggesting that a few authors uploaded their articles as separate submissions rather than updating the existing article), we extracted data for a total of 1216 articles over the last two years (figure 1A). For more details on the data collection process, see the Supplement (https://daniel1noble.github.io/ecoevo_1000/).

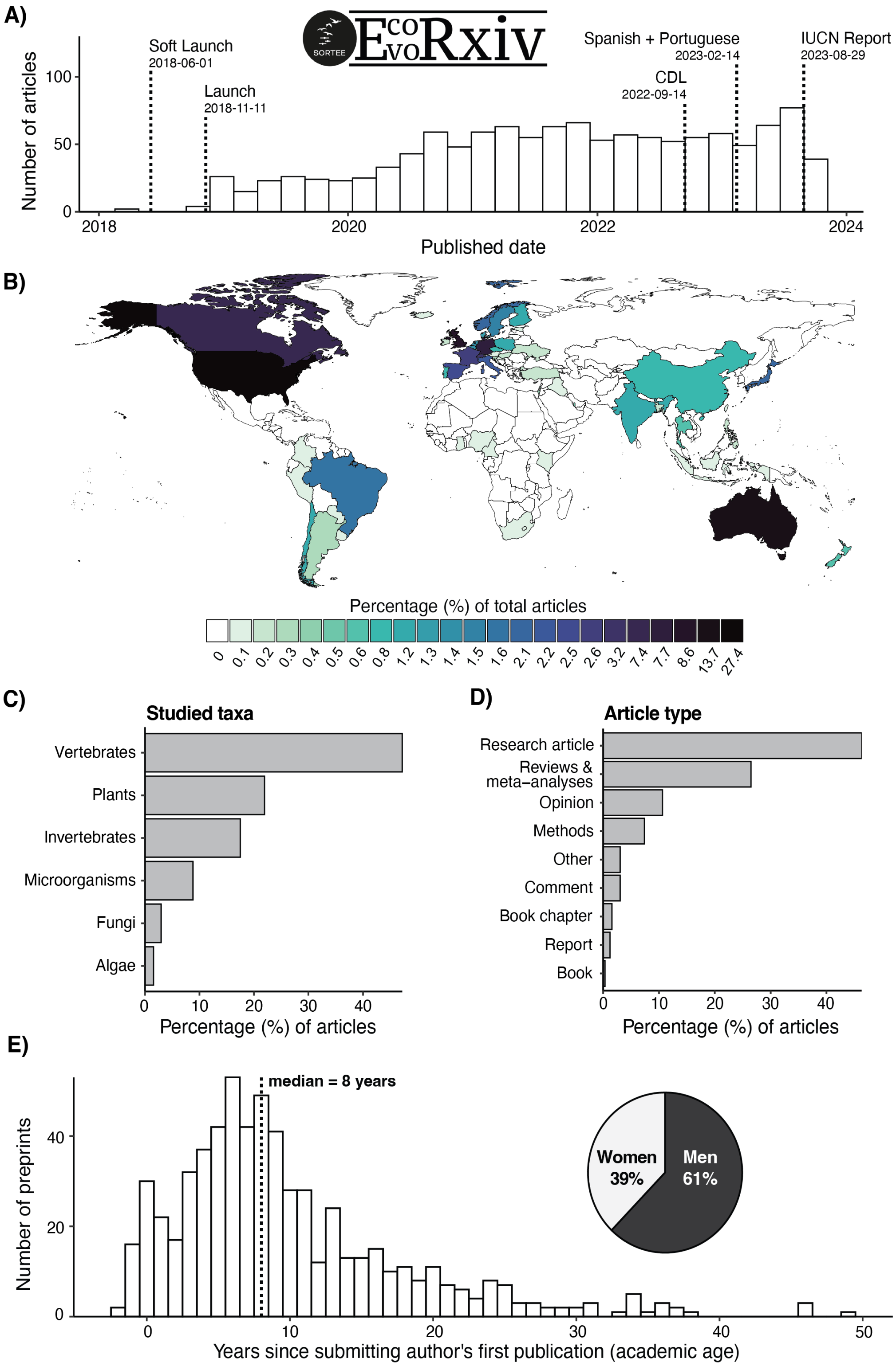

Figure 1- Summary of articles posted to *EcoEvoRxiv*. A) Number of articles (preprints and postprints) published on *EcoEvoRxiv* between 2018 and 2023. *EcoEvoRxiv* was established in June 2018 before the launch in November 2018. Notable milestones include *EcoEvoRxiv* transitioning to the California Digital Library (CDL), the acceptance of preprints and postprints in Spanish and Portuguese, and the acceptance of the first IUCN Red List Ecosystem report; B) Geographic origin of articles uploaded to *EcoEvoRxiv*, inferred from the country of affiliation of the submitting author; C) Taxa covered in the articles posted to *EcoEvoRxiv* (*n* = 1080 articles covering relevant taxa); D) Types of articles accepted on *EcoEvoRxiv* (*n* = 1216 articles); and E) Academic age of authors posting preprints to *EcoEvoRxiv* (*n* = 1135 published and

unpublished preprints) along with the gender of the submitting author. Values lower than zero are indicative of authors who uploaded preprints before their first scientific publication in a journal. Map base source: R Package maps v.3.4.2. Shapefile: Natural Earth https://www.naturalearthdata.com/about/terms-of-use/.

(a) *Overview of EcoEvoRxiv preprints (and postprints)*

EcoEvoRxiv hosts articles from authors based in 56 countries, with 90% coming from just 17 countries. North America, Australia, and European countries upload the most preprints, with many fewer coming from countries in Africa, Central America, and parts of Asia (figure 1B). Articles covered all major taxonomic groups, with the most common groups being vertebrates (47.2%), plants (21.9%), and invertebrates (17.5%) (figure 1C).

(b) *Diversifying article types on EcoEvoRxiv: overcoming the 'grey literature' problem*

Accepting a greater diversity of article types allows *EcoEvoRxiv* to help deal with the 'grey literature' problem, whereby data that are relevant for research syntheses are not published in typical peer-reviewed journals [14,15]. *EcoEvoRxiv* has made a concerted effort to diversify the types of articles hosted. This is reflected by 6.2% of of the articles being books, book chapters, reports, and other research output types, which are typically considered 'grey literature' in ecology and evolutionary biology. As a result, articles on *EcoEvoRxiv* are more diverse than those on other preprint servers which have more restrictive submission policies. For example, *bioRxiv* only accepts empirical research articles (https://www.biorxiv.org/submit-a-manuscript).

Empirical research articles are still the most common type of article submitted to *EcoEvoRxiv* (46.3%), followed by reviews and meta-analyses (26.5%) and opinion papers (10.6%) (figure 1D). Currently, *EcoEvoRxiv* does not host many reports, particularly from government or industry, but has formed fruitful partnerships with the International Union for Conservation of Nature (IUCN). For example, IUCN Red-list Ecosystem Reports are now posted to *EcoEvoRxiv* and our community has been able to work closely with the IUCN to ensure these documents meet the IUCN requirements. We authors to consider posting books, book chapters, and reports to ensure that they are openly accessible and more easily found. Accepted *EcoEvoRxiv* submissions are given a unique DOI and are indexed on Google Scholar. DOIs can be used in grant applications, CVs, and other documents to provide a link to the work.

(c) *Breaking down language barriers to scientific communication: improving diversity and data representation globally*

A significant barrier to the communication of research findings is the fact that they are primarily communicated in English [16–18]. Research communication through a single language has major consequences for the global distribution of knowledge, resulting in knowledge gaps across some of the most biodiverse and threatened regions in the world [19,20]. Such gaps also impact research syntheses and meta-analyses because they create a distorted picture of our knowledge base that can affect future research, policy development and decision-making [20–23].

EcoEvoRxiv is the only preprint server to date that breaks down language barriers to scientific communication by accepting not only English, but also Spanish, Portuguese and French language articles. *EcoEvoRxiv* plans to expand to other languages as new non-English editors for different languages become available. Such initiatives are incredibly important if we are to begin filling global gaps in scientific knowledge. However, multilingual initiatives have been slow to take off on *EcoEvoRxiv*, with only a few Spanish submissions, and a single Portuguese article posted since EcoEvoRxiv began accepting non-English articles in 2023. Part of the challenge in getting authors to submit articles in non-English languages is a lack of awareness of *EcoEvoRxiv* in non-English speaking countries, cultural differences in the perception of preprints, and a strong reliance on traditional publishing models that typically mandate publishing in English [24].

(d) *Generational and gender-based gaps in preprinting practices*

Research can take a while to be published (see below). Early and Mid-Career Researchers (EMCRs) (~10 years post-PhD) are under pressure to publish rapidly to be competitive in job applications, promotions, and obtaining grants to progress their careers [7,25]. Preprints are especially useful for EMCRs because they can achieve faster dissemination and greater visibility [4]. EMCRs may therefore be expected to make use of preprints more than colleagues at later career stages because they are more often in charge of article submission and have developed their careers in an environment where preprints are a normal part of the publication process. We collected data on the 'academic age' of submitting authors by looking at Google Scholar profiles of authors (when available) and recording their first year of publication in a peer-reviewed journal. While this is a rough estimate of career stage, there was evidence that the number of preprints posted decreases with later career stages (negative binomial glm: year slope = -0.1, SE: 0, $p < 0.001$, n = 50 years). Most preprints were submitted by authors who published their first paper in the last ~10 years (figure 1E), with the median year since first publication being 2013 (mean = 2010.7; SD = 9.9, n = 1133). These patterns support the expectation that EMCRs may use preprints to make their work more visible and disseminate their findings more quickly. However, we acknowledge that to understand the reasons why EMCRs might adopt preprint servers more readily requires community surveys, as have been done in previous studies [e.g., 26].

Gender differences in preprint use and publication outcomes have also been observed in several research fields, including ecology and evolutionary biology [27,28]. For example, gender gaps in preprint submissions were observed during COVID-19 lockdowns [29], and previous surveys have shown that female participants are less likely to suggest posting articles as preprints, suggesting gender differences in views around preprints [26]. Therefore, such discrepancies are expected to manifest in preprint use on *EcoEvoRxiv*, but it is unclear to what extent. Understanding gender publishing patterns is challenging with observational data such as ours because we cannot know the gender of authors for certain, but we can use a data-driven approach to ascertain the probability that a particular name is of a given gender (man or woman). To obtain a rough idea of an author's gender, we used the R package *gender* (v.0.6.0; [30]) to predict the most likely gender of the submitting author of a preprint. We used an algorithm to assign binary gender based on the submitting author's name. We only used the algorithm-assigned gender when the gender of a given name was identified with 95% certainty. For the remaining names, we performed manual searches to determine gender based on the pronouns and photographs from professional and personal websites. We acknowledge that our approach does not capture self-assigned and non-binary genders. As such, our assumptions about an author's gender identity may be incorrect. Our data on gender had only two missing values–one where the first name of the submitting author was missing and the other one for a collective submission. As expected, we found that women were less likely to post to *EcoEvoRxiv* compared to men (women: 38.5%, figure 1E), reinforcing existing disparities between male and female scientists. For example, studies have shown that female first authors have lower acceptance rates and are cited less (~2%) compared to males [e.g., 27].

3. Following the journey of a preprint on *EcoEvoRxiv*: from submission to publication

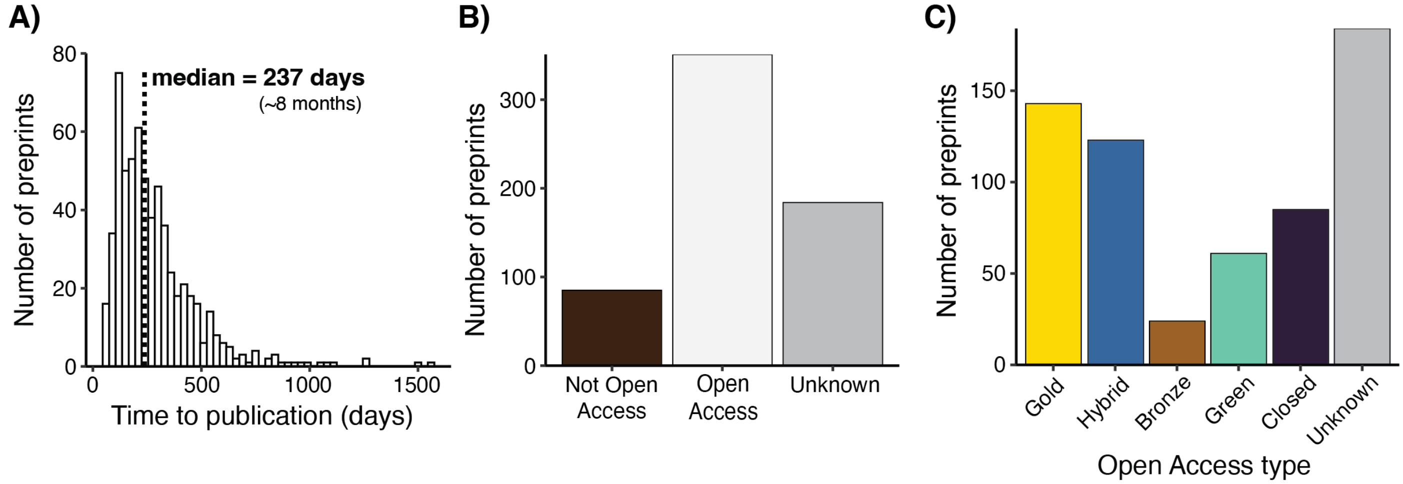

Figure 2- Summary of the publication status of preprints on *EcoEvoRxiv*. A) Time between uploading an preprint to *EcoEvoRxiv* and its publication as a peer-reviewed journal article. Preprints were considered those articles that were published in a journal a minimum of 2 months (60 days) after being posted to *EcoEvoRxiv*; B) Access status of published preprints on *EcoEvoRxiv* classified as "Open Access" or "Not Open Access". "Unknown" status is for articles who status was unclear on the Unpaywall platform; and C) Sub-types of open access status of published preprints on *EcoEvoRxiv*. Sub-type meanings are as follows: 'Green', articles published in 'toll-access journals but achieved in an open access repository; 'Bronze', articles are free to read on publishers website without a license but grants no other rights and can be delayed free-to-read; 'Hybrid', articles are free to read upon publication with an open access license; 'Gold', articles published in fully open access journals. For full details on the meaning of each category see https://support.unpaywall.org/support/solutions/articles/44001777288-what-do-the-types-of-oa-status-green-gold-hybrid-and-bronze-mean-

(a) *Science takes time, but publication could take longer*

Increased competition in science has raised the bar with respect to the amount of data required for publication [7]. This requirement is a good outcome if it results in higher-impact research that better clarifies our understanding of the natural world, but it does come at a cost for the speed of research dissemination [7]. Long publication times can adversely affect EMCRs who rely on publications for job applications, promotions, and obtaining grants. Getting research out quickly can also be critical for the development of new knowledge that can re-shape research landscapes, which was important during the COVID-19 pandemic [31]. Preprints have been proposed as a way to disseminate research more quickly as it can take a long time before results are ultimately published after formal peer review [6,7]. However, data on the time to publication is needed to quantify the real benefit of preprints in this context.

We estimated how long it takes to publish a preprint in ecology and evolution by recording the time between when a preprint was first posted on *EcoEvoRxiv*, and its final acceptance in a peer-reviewed journal. In total, 515 preprints remained unpublished (45.4%, n = 1135) at the time when these data were collected. Not all of these preprints, however, are anticipated to be published in a peer-reviewed journal (e.g., reports). Nonetheless, the median time to publication for preprints was 237 days (8 months) for all preprints that ended up being published (mean = 286.4; SD = 193.9 days) with the maximum time to publication being 1549 days or 4.2 years (figure 2A). For a full breakdown on the time to publication based on article type, refer to the Supplement (Table S1). Our results largely confirm the extended timeframes that most authors experience between writing their research papers and their publication.

(b) *Cautious 'open'-mindedness of research in preprints*

In addition to speeding up dissemination, preprints and postprints can also be a useful way to ensure that research remains open and accessible to the research community irrespective of the accessibility of the final peer-reviewed paper [6,7]. We evaluated whether preprints and postprints hosted at *EcoEvoRxiv* and that were also published in a journal were published open access. The open access status of each published article was obtained using the R package *roadoi* (v.0.7.2) to connect to the Unpaywall platform [32]. Most of the published preprints and postprints were open access [80.5% (*n* = 351 out of 436 where the status was known); figure 2B]; however, 19.5% (*n* = 85) were published behind paywalls. Published articles behind paywalls may otherwise remain inaccessible if it were not for *EcoEvoRxiv*. For preprints and postprints published in open access journals, the type of open access also varied widely (e.g., Gold, Hybrid, Green OA etc., figure 2C). Such a result may not be too surprising given that authors using preprint servers are probably already 'pro-open access', particularly given that *EcoEvoRxiv* is run by the Society for Open, Reliable, and Transparent Ecology and Evolutionary Biology (SORTEE).

Data and code sharing are also key components of open science [33]. In the spirit of 'openness', we expected data and code sharing among preprints and postprints to be greater than in many papers published in research journals [33,34]. Despite this, we found that 54.4% (*n* = 398) of articles relying on data (i.e., classified as 'research articles') on *EcoEvoRxiv* did not share data, and 58.1% (*n* = 425) did not share code.

Authors may be reluctant to share data and code for preprints because of the perceived concern that others may acquire and use their data and code before publication in a journal. Authors of 28.7% (*n* = 123) preprints did not share data at the preprint stage but ultimately did share data when the article was published, whereas authors of 35.2% (*n* = 151) never shared data and 36.1% (*n* = 155) shared data at both stages. The same was true for code. Overall, 16.8% (*n* = 72) of preprints had no open code at the preprint stage, but ultimately did share code at the published article stage. In contrast, authors of 45.2% (*n* = 194) preprints did not share code at either stage while 38% (*n* = 163) shared code at both stages. Relatively low code and data-sharing practices in our sample is consistent with analyses of sharing practices for published articles (e.g., [34]), even for journals with strict public data archiving policies [33].

4. Paving our future to open, transparent and community-driven science

Our analysis has allowed us to better understand preprinting/postprinting practices in *EcoEvoRxiv*. Overall, *EcoEvoRxiv* articles are diverse but with primary research articles on vertebrates comprising most of the articles posted. North America, Europe and Australia use *EcoEvoRxiv* the most with very few non-English language articles deposited to date. Submitting authors who were earlier in their career and more often with 'male-associated names' tended to use *EcoEvoRxiv* the most. Articles posted to *EcoEvoRxiv* tend to take up to 8 months to become published with many articles not being open access. Code and data sharing was also relatively uncommon at the preprint stage. We attempted to collect data on community discussion around preprints. However, no such data was found on preprint landing pages, likely reflecting inadequate functionality and cross-linking with sources where such discussion is occurring. Based on the insights from our analysis, we provide recommendations to authors and the scientific community on ways they can further promote open and transparent research through preprints:

- First, share your data and code at the preprint stage. Sharing data and code early can help improve the quality of research, establish precedence, and improve the transparency and computational reproducibility of scientific findings [25,35]. Reassuringly, sharing data and code is rarely associated with the 'scooping' of research findings [36]. If authors are worried about data being used unintentionally, clear information surrounding its reuse can be included in a license (see https://choosealicense.com). Data can also be archived with an embargo on its reuse [37].

- Second, take advantage of peer-reviewing services such as Peer Community In (PCI). The time between posting a preprint and publication is still quite long (~8 months). One possible explanation is that preprints are not being sent to suitable journals or are struggling to get into review, slowing down constructive feedback that can improve the quality of a paper. Using PCI circumvents editorial decisions without review, yet only 1.8% (n = 20) used PCI. Using such services will ensure that authors receive faster feedback on a paper. Ninety-three journals currently accept PCI reviews and recommendations when considering a paper for publication (https://peercommunityin.org/pci-friendly-journals/).
- Third, seek out and contribute to constructive feedback on preprints [6]. While it is clear that preprints help establish precedence and allow findings to be openly accessible, it still seems rare that constructive discussions form around preprints in an open forum (e.g. *bioRxiv* [38]). Unfortunately, the *EcoEvoRxiv* website does not provide opportunities for discussion given the limitations of the web server at this point in time. As such, we could not accurately assess how much discourse around a given preprint occurs. One way to facilitate such discussions may be to use open preprint peer-review services such as *Peer Community In* (PCI) or *PubPeer* (see also [39]) to provide feedback on preprints. Both *PCI* and *PubPeer* provide opportunities for open peer review around a preprint. Peer discussions are given a unique DOI which can then be associated with a preprint on *EcoEvoRxiv*. However, *EcoEvoRxiv* currently lacks connection to *PCI*, *PubPeer* and Altmetric data which would allow for discussion to be assimilated around a preprint in one place and make it easier for readers to follow discourse around a preprint. Clearly, as a community, we need to provide better platforms and workflows that document discussions around preprint findings. Such discussions help authors improve their work and communicate their findings more effectively (when done constructively, of course). Lack of discussion around preprints more generally might also be a function of time constraints researchers face and the lack of credit received for such community service. An important future goal of *EcoEvoRxiv* is to provide better community discussion forums and integration with existing preprint peer review services. We also need to find new ways to give credit to colleagues that contribute to community discussion.
- Finally, keep your preprints updated. While most preprints get seamlessly connected and merged with their published version, some remain 'disconnected' as separate articles. Incorrect cross-linking by indexing platforms (e.g., Google Scholar) can create confusion and lead to frustration among authors. The reasons for unmatched preprints and publications are well-understood and usually easily rectified. They often result from a mismatch between preprint and published metadata (e.g., titles and author details). For example, nearly one-third of articles changed their title from preprint submission to publication [30.5% (n = 229)]. We found that mismatched metadata almost always contributed to preprints and published articles not being matched automatically in Google Scholar. At times further manual merging by authors is needed to connect the preprint and published article (this can be done in Google Scholar). Regardless, we recommend that authors update their preprints with the publication DOI when accepted to journals, especially if their title has changed. This is very easy for authors to do on *EcoEvoRxiv* and should increase the chances that the preprint is correctly linked to the published article and citations are appropriately merged.

Despite the early successes of the new initiatives taken by *EcoEvoRxiv*, as described above, much work remains to be done to improve the understanding and use of preprints and postprints within our community. We view this perspective piece as a small step towards achieving that goal. We hope that readers are more familiar with the benefits of using community-driven preprint servers and the unique initiatives they can pursue. Community-driven preprint servers can set their own agenda and are driven by the needs and desires of the community. Supporting these initiatives should be a priority for all researchers. Volunteers at *EcoEvoRxiv* are encouraged to remain open to new and innovative ways to improve publication and open science practices. Our analysis can be used to drive changes in *EcoEvoRxiv* to make it a better platform for our community. We believe that the future of preprints is bright, and community-driven initiatives, such as *EcoEvoRxiv* will play a crucial role in the future of scientific publishing.

Data accessibility

All data and code can be found on GitHub at: https://github.com/daniel1noble/ecoevo_1000. It is also available on Zenodo, https://zenodo.org/doi/10.5281/zenodo.13841039.

Competing interests

The authors would like to acknowledge competing interests on the perspectives presented in this paper given that many (DWAN, SN, ML) are founding members of *EcoEvoRxiv* and/or are part of the *EcoEvoRxiv* committee.

Funding

DWAN would also like to thank the Australian Research Council for a Future Fellowship (FT220100276). SN and ML are supported by the Australian Research Council (ARC) Discovery Project Grants (DP210100812 and DP230101248). ARM is supported by the Natural Sciences and Engineering Research Council of Canada (RGPIN-2019-05520 and RN513790-510021).

Acknowledgements

We would like to thank the California Digital Library (CDL) and the CDL team (particularly, Alainna Wrigley, Justin Gonder, Lisa Schiff, Catherine Mitchell, Hardy Pottinger and Amanda Karby) for their support in hosting and maintaining *EcoEvoRxiv* for the Society for Open, Reliable, and Transparent Ecology and Evolutionary Biology (SORTEE). We would like to thank Gabriela Hidalgo and Daisy Larios for helping connect us with the IUCN and facilitating discussions to make *EcoEvoRxiv* a place where IUCN reports can be posted. Finally, we would also like to thank the endless number of SORTEE volunteers, and those especially on the *EcoEvoRxiv* Committee, who have helped to make *EcoEvoRxiv* a success. This paper emerged from a hackathon at the 2023 SORTEE conference, and we thank delegates who attended the session but could not be part of this paper.

References

1. Ginsparg P. 2011 ArXiv at 20. *Nature* **476**, 145–147.

2. Berg JM *et al.* 2016 Preprints for the life sciences. *Science* **352**, 899–901.

3. Colavizza G, Cadwallader L, LaFlamme M, Dozot G, Lecorney S, Rappo D, Hrynaszkiewicz I. 2024 An analysis of the effects of sharing research data, code, and preprints on citations. *arXiv preprint arXiv:2404.16171*

4. Fu DY, Hughey JJ. 2019 Releasing a preprint is associated with more attention and citations for the peer-reviewed article. *Elife* **8**, e52646.

5. Desjardins-Proulx P, White EP, Adamson JJ, Ram K, Gravel TPD. 2013 The case for open preprints in biology. *PLOS Biology* **11**, e1001563.

6. Bourne PE, Polka JK, Vale RD, Kiley R. 2017 Ten simple rules to consider regarding preprint submission. *PLOS Computational Biology* **13**, e1005473.

7. Vale RD. 2015 Accelerating scientific publication in biology. *Proceedings of the National Academy of Sciences* **112**, 13439–13446.

8. Ni R, Waltman L. 2024 To preprint or not to preprint: A global researcher survey. *Journal of the Association for Information Science and Technology*

9. Klebel T, Reichmann S, Polka J, McDowell G, Penfold N, Hindle S, Ross-Hellauer T. 2020 Peer review and preprint policies are unclear at most major journals. *PLOS ONE* **15**, e0239518.

10. Chiarelli A, Johnson R, Pinfield S, Richens E. 2019 Preprints and scholarly communication: An exploratory qualitative study of adoption, practices, drivers and barriers. *F1000 Research* **8**, 971, https://doi.org/10.12688/f1000research.19619.2.

11. Fraser N, Mayr P, Peters I. 2022 Motivations, concerns and selection biases when posting preprints: A survey of bioRxiv authors. *PLOS ONE* **17**, e0274441.

12. Guillemaud T, Facon B, Bourguet D. 2019 Peer community in: A free process for the recommendation of unpublished scientific papers based on peer review. *ELectronic PUBlishing*

13. O'Dea RE *et al.* 2021 Towards open, reliable, and transparent ecology and evolutionary biology. *BMC biology* **19**, 68.

14. Haddaway NR, Bayliss HR. 2015 Shades of grey: Two forms of grey literature important for reviews in conservation. *Biological Conservation* **191**, 827–829.

15. Haddaway NR, Bethel A, Dicks LV, Koricheva J, Macura B, Petrokofsky G, Pullin AS, Savilaakso S, Stewart GB. 2020 Eight problems with literature reviews and how to fix them. *Nature Ecology & Evolution* **4**, 1582–1589.

16. Amano T, González-Varo JP, Sutherland WJ. 2016 Languages are still a major barrier to global science. *PLOS Biology* **14**, e2000933.

17. Amano T, Sutherland WJ. 2013 Four barriers to the global understanding of biodiversity conservation: Wealth, language, geographical location and security. *Proceedings of the Royal Society B: Biological Sciences* **280**, 20122649.

18. Amano T, Rios Rojas C, Boum II Y, Calvo M, Misra BB. 2021 Ten tips for overcoming language barriers in science. *Nature Human Behaviour* **5**, 1119–1122.

19. Amano T *et al.* 2023 The role of non-english-language science in informing national biodiversity assessments. *Nature Sustainability* **6**, 845–854.

20. Zenni RD, Barlow J, Pettorelli N, Stephens P, Rader R, Siqueira T, Gordon R, Pinfield T, Nuñez MA. 2023 Multi-lingual literature searches are needed to unveil global knowledge. *Journal of Applied Ecology*

21. Hannah K, Haddaway NR, Fuller RA, Amano T. 2024 Language inclusion in ecological systematic reviews and maps: Barriers and perspectives. *Research Synthesis Methods*

22. White CR, Marshall DJ, Chown SL, Clusella-Trullas S, Portugal SJ, Franklin CE, Seebacher F. 2021 Geographical bias in physiological data limits predictions of global change impacts. *Functional Ecology* **35**, 1572–1578.

23. Konno K, Akasaka M, Koshida C, Katayama N, Osada N, Spake R, Amano T. 2020 Ignoring non-english-language studies may bias ecological meta-analyses. *Ecology and Evolution* **10**, 6373–6384.

24. Arenas-Castro H *et al.* 2024 Academic publishing requires linguistically inclusive policies. *Proceedings of the Royal Society B* **291**, 20232840.

25. Sarabipour S, Debat HJ, Emmott E, Burgess SJ, Schwessinger B, Hensel Z. 2019 On the value of preprints: An early career researcher perspective. *PLOS Biology* **17**, e3000151.

26. Fraser N, Mayr P, Peters I. 2022 Motivations, concerns and selection biases when posting preprints: A survey of bioRxiv authors. *PLoS One* **17**, e0274441.

27. Fox CW, Paine CT. 2019 Gender differences in peer review outcomes and manuscript impact at six journals of ecology and evolution. *Ecology and Evolution* **9**, 3599–3619.

28. Wehner MR, Li Y, Nead KT. 2020 Comparison of the proportions of female and male corresponding authors in preprint research repositories before and during the COVID-19 pandemic. *JAMA network open* **3**, e2020335–e2020335.

29. Ucar I, Torre M, Elías A. 2022 Mind the gender gap: COVID-19 lockdown effects on gender differences in preprint submissions. *PloS one* **17**, e0264265.

30. Mullen L. 2021 Predict gender from names using historical data. R package version 0.6.0.

31. Älgå A, Eriksson O, Nordberg M. 2021 The development of preprints during the COVID-19 pandemic. *Journal of Internal Medicine* **290**, 480.

32. Jahn N. 2024 Roadoi: Find free versions of scholarly publications via unpaywall. R package version 0.7.2. Https://github.com/ropensci/roadoi/, https://docs.ropensci.org/roadoi/.

33. Roche DG, Kruuk LE, Lanfear R, Binning SA. 2015 Public data archiving in ecology and evolution: How well are we doing? *PLOS Biology* **13**, e1002295.

34. O'Dea RE *et al.* 2021 Preferred reporting items for systematic reviews and meta-analyses in ecology and evolutionary biology: A PRISMA extension. *Biological Reviews* **96**, 1695–1722.

35. Gomes DG *et al.* 2022 Why don't we share data and code? Perceived barriers and benefits to public archiving practices. *Proceedings of the Royal Society B* **289**, 20221113.

36. Soeharjono S, Roche DG. 2021 Reported individual costs and benefits of sharing open data among canadian academic faculty in ecology and evolution. *BioScience* **71**, 750–756.

37. Roche DG, Lanfear R, Binning SA, Haff TM, Schwanz LE, Cain KE, Kokko H, Jennions MD, Kruuk LE. 2014 Troubleshooting public data archiving: Suggestions to increase participation. *PLoS biology* **12**, e1001779.

38. Anderson KR. 2020 bioRxiv: Trends and analysis of five years of preprints. *Learned Publishing* **33**, 104–109.

39. Avissar-Whiting M *et al.* 2024 Recommendations for accelerating open preprint peer review to improve the culture of science. *PLOS Biology* **22**, e3002502.