



**HAL**  
open science

## The promise of community-driven preprints in ecology and evolution

Daniel Noble, Zoe Xirocostas, Nicholas Wu, April Martinig, Rafaela Almeida, Kevin Bairos-Novak, Heikel Balti, Michael Bertram, Louis Bliard, Jack Brand, et al.

### ► To cite this version:

Daniel Noble, Zoe Xirocostas, Nicholas Wu, April Martinig, Rafaela Almeida, et al.. The promise of community-driven preprints in ecology and evolution. 2024. hal-04660631

**HAL Id: hal-04660631**

**<https://hal.science/hal-04660631v1>**

Preprint submitted on 24 Jul 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - ShareAlike 4.0 International License

# The promise of community-driven preprints in ecology and evolution

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

## Affiliations

<sup>1</sup>Division of Ecology and Evolution, Research School of Biology, The Australian National University, Canberra, ACT 2600, Australia

<sup>2</sup> School of Life Sciences, Faculty of Science, University of Technology Sydney, NSW 2007, Australia

<sup>3</sup>Hawkesbury Institute for the Environment, Western Sydney University, Richmond, NSW 2753, Australia

<sup>4</sup>Evolution and Ecology Research Centre, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW 2052, Australia

<sup>5</sup>Laboratory of Freshwater Ecology, Evolution and Conservation, KU Leuven, Belgium

<sup>6</sup>School of the Environment, The University of Queensland, Brisbane 4072, Australia

<sup>7</sup> Chrono-environnement UMR 6249, CNRS, Université Bourgogne Franche-Comté, F-25000, Besançon, France

<sup>8</sup>Department of Wildlife, Fish, and Environmental Studies, Swedish University of Agricultural Sciences, Umeå 907 36, Sweden

<sup>9</sup>Department of Zoology, Stockholm University, Stockholm 114 18, Sweden

<sup>10</sup>School of Biological Sciences, Monash University, Melbourne 3800, Australia

<sup>11</sup> Department of Evolutionary Biology and Environmental Studies, University of Zurich, Switzerland

<sup>12</sup> Department of Wildlife, Fish, and Environmental Studies, Swedish University of Agricultural Sciences, Umeå, Sweden

<sup>13</sup> School of the Environment, The University of Queensland, Brisbane 4072, Australia

<sup>14</sup> Swiss Ornithological Institute, Sempach, Switzerland

<sup>15</sup> K. Lisa Yang Center for Conservation Bioacoustics, Cornell Lab of Ornithology, Cornell University, Ithaca, NY, USA

<sup>16</sup> Theoretical and Experimental Ecology Station (SETE), CNRS, France

<sup>17</sup> Biodiversity Unit, University of Turku, 20014 Turku, Finland

<sup>18</sup> Fenner School of Environment and Society, Australian National University, Canberra ACT

<sup>19</sup> Ruđer Bošković Institute, Zagreb, Croatia

<sup>20</sup> Department of Ecology and Evolutionary Biology, University of California, Los Angeles

<sup>21</sup> Department of Biology, Case Western Reserve University, Cleveland, OH 44106, USA

<sup>22</sup> Department of Science Laboratory Technology, University of Jos, Jos, Nigeria

<sup>23</sup> Department of Ecology, University of São Paulo, Brazil

- 47 <sup>24</sup> Cluster of Excellence “Science of Intelligence”, Technical University of Berlin, Berlin, Germany  
48 <sup>25</sup> Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin,  
49 Germany  
50 <sup>26</sup> Department of Evolutionary Biology, Bielefeld University, Germany  
51 <sup>27</sup> Institute of Environmental Sciences, Jagiellonian University, Krakow, Poland  
52 <sup>28</sup> Department of Environmental Systems Science, ETH Zurich, Switzerland  
53 <sup>29</sup> Institute of Plant Sciences and Oeschger Centre for Climate Change Research, University of Bern  
54 <sup>30</sup> Department of Biology, University of Oxford, United Kingdom  
55 <sup>31</sup> Norwegian Institute for Nature Research, Trondheim, Norway  
56 <sup>32</sup> Port Perry, Ontario, Canada  
57 <sup>33</sup> Norwegian Biodiversity Information Centre, Trondheim, Norway  
58 <sup>34</sup> School of Biodiversity, One Health and Veterinary Medicine, University of Glasgow, UK  
59 <sup>35</sup> European Centre for the Environment and Human Health, University of Exeter Medical School, Penryn,  
60 United Kingdom  
61 <sup>36</sup> Ecology department, Swedish University of Agricultural sciences, Sweden  
62 <sup>37</sup> Department of Animal Ecology and Tropical Biology, Biocenter, University of Würzburg, Würzburg,  
63 Germany  
64 <sup>38</sup> Statistical Ecotoxicology, Bayreuth Center of Ecology and Environmental Research (BayCEER),  
65 University of Bayreuth, Germany  
66 <sup>39</sup> School of Life and Environmental Sciences, Deakin University, Waurin Ponds, 3216, Australia  
67 <sup>40</sup> Department of Biology, Faculty of Science, Hokkaido University, Japan  
68 <sup>41</sup> Department of Biology, Lund University, Sweden  
69 <sup>42</sup> Centre for Brain Research, University of Auckland, New Zealand  
70 <sup>43</sup> School of Agriculture, Food and Ecosystem Sciences, University of Melbourne  
71 <sup>44</sup> Department of Biology, University of Oxford  
72 <sup>45</sup> Theoretical and Experimental Ecology Station (SETE), CNRS, France  
73 <sup>46</sup> Department of Geosciences, Florida Atlantic University, Boca Raton, FL, USA  
74 <sup>47</sup> Institute of Ecology and Evolution, University of Edinburgh, UK  
75 <sup>48</sup> Department of Biology, Carleton University, Canada  
76 <sup>49</sup> Université Paris-Saclay, INRAE, AgroParisTech, UMR EcoSys, Palaiseau, France  
77 <sup>50</sup> Fisheries Department, Faculty of Natural Resources, University of Guilan, Swomeh Sara, Iran  
78 <sup>51</sup> Center for Advanced Systems Understanding (CASUS), Helmholtz-Zentrum Dresden-Rossendorf  
79 (HZDR), Görlitz, Germany  
80 <sup>52</sup> Institute of Environmental Change & Society, University of Regina  
81 <sup>53</sup> Division of Behavioural Ecology, University of Bern, Switzerland  
82 <sup>54</sup> Department of Computational Landscape Ecology, Helmholtz Center for Environmental Research - UFZ,  
83 Leipzig, Germany  
84 <sup>55</sup> TRANSfarm - Science, Engineering, & Technology Group, KU Leuven, Lovenjoel 3360, Belgium  
85 <sup>56</sup> The Arctic University Museum of Norway, The Arctic University of Norway (UiT), Norway  
86 <sup>57</sup> Helsinki Institute of Life Sciences, Helsinki University, Finland  
87 <sup>58</sup> Department of Ecology, Radboud University Nijmegen, Nijmegen  
88 <sup>59</sup> Netherlands Institute of Ecology (NIOO-KNAW), Wageningen, Netherlands  
89 <sup>60</sup> School of BioSciences, The University of Melbourne, Victoria 3010, Australia.  
90 <sup>61</sup> Department of Biological Sciences, University of Alberta, CW 405, Biological Sciences Building,  
91 Edmonton, AB T6G 2E9, Canada

92  
93 corresponding author, [daniel.noble@anu.edu.au](mailto:daniel.noble@anu.edu.au)

94 \* equal contribution

95 <sup>5-60</sup>, Authors listed in alphabetical order

96

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

## 112 **1. Introduction**

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

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

131 *EcoEvoRxiv* is one of the few community-driven preprint servers that has paved the way for new initiatives,  
132 by accepting multilingual preprints, registered reports, and non-traditional research reports. *EcoEvoRxiv* also  
133 promotes peer review and community discussion in the hopes of improving the quality of preprints and  
134 speeding up their peer-reviewed publication (e.g., Peer Community In – PCI [12]). Peer Community In  
135 allows for fast, constructive peer review around a preprint with peer reviews being transparent and published  
136 online [12].

137 Here, we explore the first preprints uploaded to *EcoEvoRxiv* to characterise preprint practices in ecology and  
138 evolution. We aim to understand: 1) in what countries authors who use *EcoEvoRxiv* are located; 2) the  
139 taxonomic diversity study systems across preprints; 3) whether preprint server use depends on career stage  
140 and gender; 4) the extent to which authors make use of preprint servers for reports and community-driven  
141 peer review; 5) the extent to which data and code are shared in preprints; and 6) how many preprints remain

142 unpublished, and for those that are published, how long it took for them to become published. In the process,  
143 we also provide a summary of what makes *EcoEvoRxiv* distinct from other preprint servers to help further  
144 clarify the benefits of using community-driven preprint servers to disseminate research findings.

## 145 **2. Getting to know your *EcoEvoRxiv* preprint server**

146 *EcoEvoRxiv* is run by the Society for Open, Reliable, and Transparent Ecology and Evolutionary Biology  
147 (SORTEE)[13]. Originally launched in 2018 on the Center for Open Science preprint platform, *EcoEvoRxiv*  
148 has become a popular preprint server for ecologists and evolutionary biologists. The server has since been  
149 adopted by the California Digital Library (CDL). Editors are ecologists and evolutionary biologists from  
150 across the globe who volunteer their time to screen papers and push new initiatives in the preprint space.  
151 *EcoEvoRxiv* allows authors to post both preprints and postprints (also known as author-accepted  
152 manuscripts). While preprints are versions of manuscripts posted by authors before peer-review, postprints  
153 are versions of peer-reviewed and accepted articles but without typesetting and formatting by a journal. The  
154 main reason for publishing postprints on a preprint server is to ensure published articles are openly accessible  
155 to everyone without a paywall (i.e., green open access). Postprints can be published anytime, provided that  
156 journals allow it (which most do; see <https://www.sherpa.ac.uk/romeo/>).

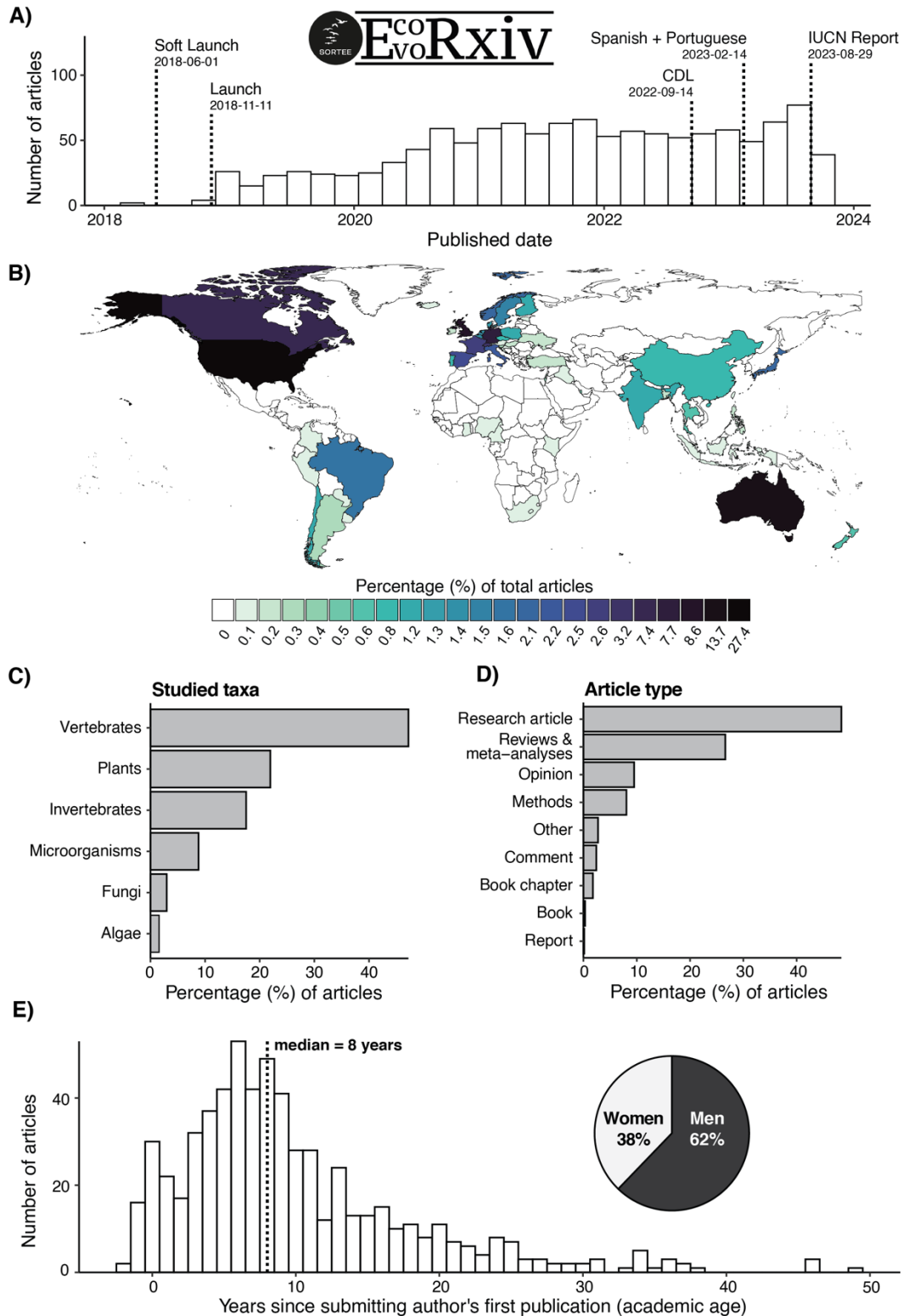


Figure 1- Summary of articles posted to *EcoEvoRxiv*. A) Number of articles (preprints and postprints) published on *EcoEvoRxiv* between 2018 and 2023. *EcoEvoRxiv* was established in June 2018 before the launch in November 2018. Notable milestones include *EcoEvoRxiv* transitioning to the California 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 the preprints and postprints uploaded to *EcoEvoRxiv*, inferred from the country of affiliation of the submitting author; C) Taxa used/covered in the articles posted to *EcoEvoRxiv* (n = 1080 articles); D) Types of preprints accepted on *EcoEvoRxiv* (n = 620 articles). E) Academic age of authors posting to *EcoEvoRxiv* 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.natureearthdata.com/about/terms-of-use/>.

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

158 To better understand preprint (and postprint) use on *EcoEvoRxiv*, we downloaded metadata on the articles  
159 available on *EcoEvoRxiv* as of 2023-09-30 (see Supplement for more details on methods). We consider both  
160 preprints and postprints as ‘articles’. After removing five duplicate titles – suggesting that a few authors  
161 created multiple submissions for the same preprint rather than updating the existing submission – we had  
162 data for a total of 1216 articles with ~55–60 preprints published approximately monthly in the last two years  
163 (figure 1A). *EcoEvoRxiv* hosts articles from authors based in 56 countries, with 90% of the articles coming  
164 from just 17 countries. North America, Australia, and European countries upload the most preprints, with  
165 many fewer coming from countries in Africa, Central America, and parts of Asia (figure 1B). Articles  
166 covered all major taxonomic groups, with the most common being vertebrates (47.2%), plants (21.9%), and  
167 invertebrates (17.5%) (figure 1C).

## 168 **(b) Diversifying article types on *EcoEvoRxiv*: overcoming the ‘grey literature’ problem**

169 Accepting a greater diversity of article types allows *EcoEvoRxiv* to help deal with the ‘grey literature’  
170 problem, whereby data that are relevant for research syntheses are not published in typical peer-reviewed  
171 journals [14,15]. *EcoEvoRxiv* has made a concerted effort to diversify the types of articles accepted. This is  
172 reflected by 5% of the articles on *EcoEvoRxiv* being books, book chapters, reports, and other research output  
173 types, which are typically considered ‘grey literature’ in ecology and evolutionary biology. As a result,  
174 articles on *EcoEvoRxiv* are more diverse than those on preprint servers which have more restrictive  
175 submission policies. For example, *bioRxiv* only accepts research articles ([https://www.biorxiv.org/submit-a-](https://www.biorxiv.org/submit-a-manuscript)  
176 [manuscript](https://www.biorxiv.org/submit-a-manuscript)).

177 Typical research articles are still the most common type of preprint on *EcoEvoRxiv* (48.4%), followed by  
178 reviews and meta-analyses (26.6%) and opinion papers (9.5%) (figure 1D). Currently, *EcoEvoRxiv* does not  
179 host many reports, particularly from government or industry, but has formed fruitful partnerships with the  
180 International Union for Conservation of Nature (IUCN). For example, IUCN Red-list Ecosystem Reports are  
181 now posted to *EcoEvoRxiv* and our community has been able to work closely with the IUCN to ensure these  
182 documents meet the IUCN requirements. We would encourage authors to consider posting books, book  
183 chapters, and reports to ensure that they are openly accessible and more easily found. *EcoEvoRxiv* articles are  
184 given a unique DOI and are indexed on Google Scholar.

## 185 **(c) Breaking down language barriers to scientific communication: improving diversity and** 186 **data representation globally**

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

193 *EcoEvoRxiv* is the only preprint server to date that breaks down language barriers to scientific  
194 communication by accepting not only English, but also Spanish, Portuguese and French language articles.  
195 *EcoEvoRxiv* plans to expand to other languages as new non-English editors become available. Such  
196 initiatives are incredibly important if we are to begin filling global voids of scientific knowledge. However,  
197 multilingual initiatives have been slow to take off on *EcoEvoRxiv*, with only a few Spanish articles, and a

198 single Portuguese article, posted since starting to accept non-English articles in 2023. Part of the challenge in  
199 getting authors to submit non-English articles is the lack of awareness of *EcoEvoRxiv* in non-English  
200 speaking countries, cultural differences in the perception of preprints, and a strong reliance on traditional  
201 publishing models that typically mandate publishing in English [24].

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

203 Research papers can take a while to be published (see below). However, Early and Mid-Career Researchers  
204 (EMCRs) (~10 years post-PhD) are under pressure to publish rapidly to be competitive in job applications,  
205 promotions, and obtaining grants to progress their careers [7,25]. Preprints are one way EMCRs can achieve  
206 faster dissemination and greater visibility [4]. As such, EMCRs may be expected to make use of preprints  
207 more than colleagues at later career stages. We collected data on the ‘academic age’ of submitting authors by  
208 looking at Google Scholar profiles of authors (when available) and recording their first year of publication in  
209 a peer-reviewed journal. While this is a rough estimate of career stage, there was evidence that the number of  
210 preprints posted decreases with later career stages (negative binomial glm: year slope = -0.1, SE: 0,  $p <$   
211 0.001,  $n = 42$  years). Most preprints were submitted by authors who published their first paper in the last ~10  
212 years (figure 1E), with the median year since first publication being 2013 (mean = 2010.6; SD = 9.9,  $n =$   
213 1214). These patterns support the expectation that EMCRs may use preprints to make their work more visible  
214 and disseminate their findings more quickly. However, we acknowledge that validating this conclusion does  
215 require more rigorous experimental approaches.

216 Gender differences in preprint use and publication outcomes have also been observed in several research  
217 fields, including ecology and evolutionary biology [26,27]. Therefore, such discrepancies are expected to  
218 manifest in preprint use on *EcoEvoRxiv*, but it is unclear to what extent. Understanding gender publishing  
219 patterns is challenging with observational data such as ours because we cannot know the gender of authors  
220 for certain, but we can use a data-driven approach to ascertain the probability that a particular name is of a  
221 given gender (man or woman). To obtain a rough idea of an author’s gender, we used the R package *gender*  
222 (v.0.6.0; [28]) to predict the most likely gender of the submitting author of a preprint. We only used  
223 algorithm-assigned names where the gender was identified with 95% certainty. For the remaining names, we  
224 performed manual searches to determine gender based on the pronouns and photographs from professional  
225 and personal websites. We acknowledge that our approach does not capture self-assigned and non-binary  
226 genders. As such, our assumptions about an author’s gender identity may be incorrect. Our data on gender  
227 had only two missing values—one where the first name of the submitting author was missing and the other  
228 one for a collective submission. As expected, we found that women were less likely to publish on  
229 *EcoEvoRxiv* compared to men (women: 37.8%), which may reflect the broader publishing disparities  
230 between male and female scientists [26].



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

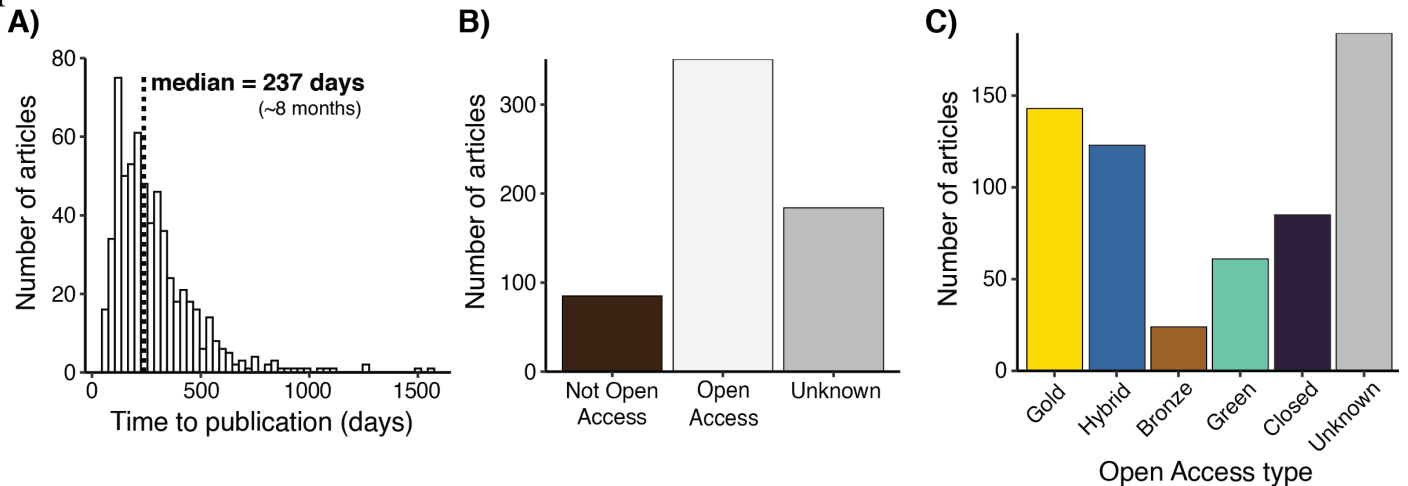


Figure 2- Summary of the publication status of articles on *EcoEvoRxiv*. A) Time between uploading an preprint to *EcoEvoRxiv* and its publication as a peer-reviewed journal article. Articles that were published within 2 months (60 days) of being submitted are considered postprints and are not included in this figure. B) Access status of articles linked to *EcoEvoRxiv* preprints classified as “Open Access” or “Not Open Access”. “Unknown” status is for articles not explicitly identified as being open access or not on the Unpaywall platform. C) Sub-types of open access status of articles linked to *EcoEvoRxiv* preprints. 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]. Preprints have been proposed as a way to disseminate research more quickly as it can take a long time before results are ultimately published [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 research paper in ecology and evolution by recording the time between when an article was first posted on *EcoEvoRxiv*, and its final acceptance in a peer-reviewed journal. In total, 383 papers remained unpublished (31.5%,  $n = 1216$ ) at the time when these data were collected. Not all of these papers, 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) (mean = 286.4; SD = 193.9 days) with the maximum time to publication being 1549 days or 4.2 years (figure 2A). 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 articles 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 [29]. Most of the published articles

254 were open access [80.5% ( $n = 351$  out of 436 where the status was known); figure 2A]; however, 19.5% ( $n =$   
255 85) were published behind paywalls. For articles published in open access journals, the type of open access  
256 also varied widely (e.g., Gold, Hybrid, Green OA etc., figure 2B).

257 Data and code sharing are also key components of open science [30]. In the spirit of ‘openness’, we expected  
258 data and code sharing among preprints to be greater than in many papers published in research journals  
259 [30,31]. Despite this, we found that 54.1% ( $n = 232$ ) of articles relying on data on *EcoEvoRxiv* did not share  
260 data, and 56.9% ( $n = 246$ ) did not share code (counting only data-based articles, i.e., excluding reviews,  
261 commentaries or theoretical works).

262 Authors may be reluctant to share data and code for preprints because of the perceived concern that others  
263 may acquire and use their data and code before publication in a journal. Authors of 28.7% ( $n = 123$ ) of  
264 articles that did not share data at the preprint stage did ultimately share data when the article was published;  
265 whereas authors of 35.2% ( $n = 151$ ) never shared data. However, 36.1% ( $n = 155$ ) shared data at both stages.  
266 The same was true for code. Overall, 16.8% ( $n = 72$ ) preprints had no open code at the preprint stage but did  
267 at the published article stage and authors of 45.2% ( $n = 194$ ) preprints did not share code at either stage.  
268 However, 38% ( $n = 163$ ) shared code at both stages. Relatively low code and data-sharing practices in our  
269 sample is consistent with analyses of sharing practices for published articles (e.g., [31]), even for journals  
270 with strict public data archiving policies [30].

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

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

- 283 • First, share your data and code at the preprint stage. Sharing data and code early can help improve the  
284 quality of research, establish precedence, and improve the transparency and computational  
285 reproducibility of scientific findings [25]. Reassuringly, sharing data and code is rarely associated  
286 with the ‘scooping’ of research findings [32]. If authors are worried about data being used  
287 unintentionally, clear information surrounding its reuse can be included in a license (see  
288 <https://choosealicense.com>). Data can also be archived with an embargo on its reuse [33].
- 289 • Second, take advantage of peer-reviewing services such as Peer Community In (PCI). The time  
290 between posting a preprint and publication is still quite long (~8 months). One possible explanation is  
291 that preprints are not being sent to suitable journals or are struggling to get into review, slowing down  
292 constructive feedback that can improve the quality of a paper. Using PCI circumvents editorial  
293 decisions without review, yet only 1.6% ( $n = 10$ ) used PCI. Using such services will ensure that  
294 authors receive faster feedback on a paper. Ninety-three journals currently accept PCI reviews and  
295 recommendations when considering a paper for publication ([https://peercommunityin.org/pci-](https://peercommunityin.org/pci-friendly-journals/)  
296 [friendly-journals/](https://peercommunityin.org/pci-friendly-journals/)).
- 297 • Third, seek out and contribute to constructive feedback on preprints [6]. While it is clear that  
298 preprints help establish precedence and allow findings to be openly accessible, it still seems rare that  
299 constructive discussions form around preprints in an open forum (e.g. bioRxiv [34]). Unfortunately,

300 the *EcoEvoRxiv* website does not provide opportunities for discussion given the limitations of the web  
301 server at this point in time. As such, we could not accurately assess how much discourse around a  
302 given preprint occurs. Clearly, as a community, we need to provide better platforms that document  
303 discussions around preprint findings. Such discussions help authors improve their work and  
304 communicate their findings more effectively (when done constructively, of course). One way to  
305 facilitate such discussions may be to use open preprint peer-review services such as *Peer Community*  
306 *In* (PCI) or *PubPeer* (see also [35]) to provide feedback on preprints.

- 307 • Finally, keep your preprints updated. While most preprints get seamlessly connected and merged with  
308 their published version, some remain ‘disconnected’ as separate articles. Incorrect cross-linking by  
309 indexing platforms (e.g., Google Scholar) can create confusion and lead to frustration among authors.  
310 Thankfully, the reasons for unmatched preprints and publications are well-understood and easily  
311 rectified. They often result from a mismatch between preprint and published metadata (e.g., titles and  
312 author details). For example, nearly one-third of articles changed their title from preprint submission  
313 to publication [32.3% ( $n = 199$ )]. We found that mismatched metadata almost always contributed to  
314 preprints and published articles not being matched automatically in Google Scholar. We recommend  
315 that authors update their preprints with the publication DOI when accepted to journals, especially if  
316 their title has changed. This is very easy for authors to do on *EcoEvoRxiv* and ensures that the preprint  
317 is correctly linked to the published article and citations are appropriately merged. Adding DOIs will  
318 also create a link between the final preprint and the published paper so that, no matter the open access  
319 status of the publication, the research findings will remain openly accessible, saving costs for authors  
320 (i.e., by not having to pay open access fees) and improving the visibility and use of research.

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

## 330 **Data accessibility**

331 All data and code can be found on GitHub at: [https://github.com/daniel1noble/econvo\\_1000](https://github.com/daniel1noble/econvo_1000)

## 332 **Competing interests**

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

## 336 **Funding**

337 DWAN would also like to thank the Australian Research Council for a Future Fellowship (FT220100276).  
338 SN and ML are supported by the Australian Research Council (ARC) Discovery Project Grants  
339 (DP210100812 and DP230101248).

## 340 Acknowledgements

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

## 350 References

- 351 1. Ginsparg P. 2011 ArXiv at 20. *Nature* **476**, 145–147.
- 352 2. Berg JM *et al.* 2016 Preprints for the life sciences. *Science* **352**, 899–901.
- 353 3. Colavizza G, Cadwallader L, LaFlamme M, Dozot G, Lecorney S, Rappo D, Hrynaszkiewicz I. 2024  
354 An analysis of the effects of sharing research data, code, and preprints on citations. *arXiv preprint*  
355 *arXiv:2404.16171*
- 356 4. Fu DY, Hughey JJ. 2019 Releasing a preprint is associated with more attention and citations for the  
357 peer-reviewed article. *Elife* **8**, e52646.
- 358 5. Desjardins-Proulx P, White EP, Adamson JJ, Ram K, Gravel TPD. 2013 The case for open preprints  
359 in biology. *PLOS Biology* **11**, e1001563.
- 360 6. Bourne PE, Polka JK, Vale RD, Kiley R. 2017 Ten simple rules to consider regarding preprint  
361 submission. *PLOS Computational Biology* **13**, e1005473.
- 362 7. Vale RD. 2015 Accelerating scientific publication in biology. *Proceedings of the National Academy*  
363 *of Sciences* **112**, 13439–13446.
- 364 8. Ni R, Waltman L. 2024 To preprint or not to preprint: A global researcher survey. *Journal of the*  
365 *Association for Information Science and Technology*
- 366 9. Klebel T, Reichmann S, Polka J, McDowell G, Penfold N, Hindle S, Ross-Hellauer T. 2020 Peer  
367 review and preprint policies are unclear at most major journals. *PLOS ONE* **15**, e0239518.
- 368 10. Chiarelli A, Johnson R, Pinfield S, Richens E. 2019 Preprints and scholarly communication: An  
369 exploratory qualitative study of adoption, practices, drivers and barriers. *F1000 Research* **8**, 971,  
370 <https://doi.org/10.12688/f1000research.19619.2>.
- 371 11. Fraser N, Mayr P, Peters I. 2022 Motivations, concerns and selection biases when posting preprints:  
372 A survey of bioRxiv authors. *PLOS ONE* **17**, e0274441.
- 373 12. Guillemaud T, Facon B, Bourguet D. 2019 Peer community in: A free process for the  
374 recommendation of unpublished scientific papers based on peer review. *ELectronic PUBlishing*
- 375 13. O’Dea RE *et al.* 2021 Towards open, reliable, and transparent ecology and evolutionary biology.  
376 *BMC biology* **19**, 68.

- 377 14. Haddaway NR, Bayliss HR. 2015 Shades of grey: Two forms of grey literature important for reviews  
378 in conservation. *Biological Conservation* **191**, 827–829.
- 379 15. Haddaway NR, Bethel A, Dicks LV, Koricheva J, Macura B, Petrokofsky G, Pullin AS, Savilaakso S,  
380 Stewart GB. 2020 Eight problems with literature reviews and how to fix them. *Nature Ecology &*  
381 *Evolution* **4**, 1582–1589.
- 382 16. Amano T, González-Varo JP, Sutherland WJ. 2016 Languages are still a major barrier to global  
383 science. *PLOS Biology* **14**, e2000933.
- 384 17. Amano T, Sutherland WJ. 2013 Four barriers to the global understanding of biodiversity  
385 conservation: Wealth, language, geographical location and security. *Proceedings of the Royal Society B:*  
386 *Biological Sciences* **280**, 20122649.
- 387 18. Amano T, Rios Rojas C, Boum II Y, Calvo M, Misra BB. 2021 Ten tips for overcoming language  
388 barriers in science. *Nature Human Behaviour* **5**, 1119–1122.
- 389 19. Amano T *et al.* 2023 The role of non-english-language science in informing national biodiversity  
390 assessments. *Nature Sustainability* **6**, 845–854.
- 391 20. Zenni RD, Barlow J, Pettorelli N, Stephens P, Rader R, Siqueira T, Gordon R, Pinfield T, Nuñez MA.  
392 2023 Multi-lingual literature searches are needed to unveil global knowledge. *Journal of Applied Ecology*
- 393 21. Hannah K, Haddaway NR, Fuller RA, Amano T. 2024 Language inclusion in ecological systematic  
394 reviews and maps: Barriers and perspectives. *Research Synthesis Methods*
- 395 22. White CR, Marshall DJ, Chown SL, Clusella-Trullas S, Portugal SJ, Franklin CE, Seebacher F. 2021  
396 Geographical bias in physiological data limits predictions of global change impacts. *Functional Ecology*  
397 **35**, 1572–1578.
- 398 23. Konno K, Akasaka M, Koshida C, Katayama N, Osada N, Spake R, Amano T. 2020 Ignoring non-  
399 english-language studies may bias ecological meta-analyses. *Ecology and Evolution* **10**, 6373–6384.
- 400 24. Arenas-Castro H *et al.* 2024 Academic publishing requires linguistically inclusive policies.  
401 *Proceedings of the Royal Society B* **291**, 20232840.
- 402 25. Sarabipour S, Debat HJ, Emmott E, Burgess SJ, Schwessinger B, Hensel Z. 2019 On the value of  
403 preprints: An early career researcher perspective. *PLOS Biology* **17**, e3000151.
- 404 26. Fox CW, Paine CT. 2019 Gender differences in peer review outcomes and manuscript impact at six  
405 journals of ecology and evolution. *Ecology and Evolution* **9**, 3599–3619.
- 406 27. Wehner MR, Li Y, Nead KT. 2020 Comparison of the proportions of female and male corresponding  
407 authors in preprint research repositories before and during the COVID-19 pandemic. *JAMA network open*  
408 **3**, e2020335–e2020335.
- 409 28. Mullen L. 2021 Predict gender from names using historical data. R package version 0.6.0.
- 410 29. Jahn N. 2024 Roadoi: Find free versions of scholarly publications via unpaywall. R package version  
411 0.7.2. <https://github.com/ropensci/roadoi/>, <https://docs.ropensci.org/roadoi/>.
- 412 30. Roche DG, Kruuk LE, Lanfear R, Binning SA. 2015 Public data archiving in ecology and evolution:  
413 How well are we doing? *PLOS Biology* **13**, e1002295.

- 414 31. O’Dea RE *et al.* 2021 Preferred reporting items for systematic reviews and meta-analyses in ecology  
415 and evolutionary biology: A PRISMA extension. *Biological Reviews* **96**, 1695–1722.
- 416 32. Soeharjono S, Roche DG. 2021 Reported individual costs and benefits of sharing open data among  
417 canadian academic faculty in ecology and evolution. *BioScience* **71**, 750–756.
- 418 33. Roche DG, Lanfear R, Binning SA, Haff TM, Schwanz LE, Cain KE, Kokko H, Jennions MD, Kruuk  
419 LE. 2014 Troubleshooting public data archiving: Suggestions to increase participation. *PLoS biology* **12**,  
420 e1001779.
- 421 34. Anderson KR. 2020 bioRxiv: Trends and analysis of five years of preprints. *Learned Publishing* **33**,  
422 104–109.
- 423 35. Avissar-Whiting M *et al.* 2024 Recommendations for accelerating open preprint peer review to  
424 improve the culture of science. *PLOS Biology* **22**, e3002502.
- 425