

# 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 Introduction

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

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

115 *EcoEvoRxiv* is one of the few community-driven preprint servers that has paved the way for new initiatives,  
 116 by accepting multilingual preprints, registered reports, and non-traditional research reports. *EcoEvoRxiv* also  
 117 promotes peer review and community discussion in the hopes of improving the quality of preprints and  
 118 speeding up their peer-reviewed publication (e.g., Peer Community In – PCI<sup>12</sup>). Peer Community In allows  
 119 for fast, constructive peer review around a preprint with peer reviews being transparent and published  
 120 online<sup>12</sup>.

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

## 129 Getting to know your *EcoEvoRxiv* preprint server

130 *EcoEvoRxiv* is run by the Society for Open, Reliable, and Transparent Ecology and Evolutionary Biology  
 131 (SORTEE)<sup>13</sup>. Originally launched in 2018 on the Center for Open Science preprint platform, *EcoEvoRxiv* has  
 132 become a popular preprint server for ecologists and evolutionary biologists. The server has since been  
 133 adopted by the California Digital Library (CDL). Editors are ecologists and evolutionary biologists from  
 134 across the globe who volunteer their time to screen papers and push new initiatives in the preprint space.  
 135 *EcoEvoRxiv* allows authors to post both preprints and postprints (also known as author-accepted  
 136 manuscripts). While preprints are versions of manuscripts posted by authors before peer-review, postprints  
 137 are versions of peer-reviewed and accepted articles but without typesetting and formatting by a journal. The  
 138 main reason for publishing postprints on a preprint server is to ensure published articles are openly accessible

139 to everyone without a paywall (i.e., green open access). Postprints can be published anytime, provided that  
 140 journals allow it (which most do; see <https://www.sherpa.ac.uk/romeo/>).

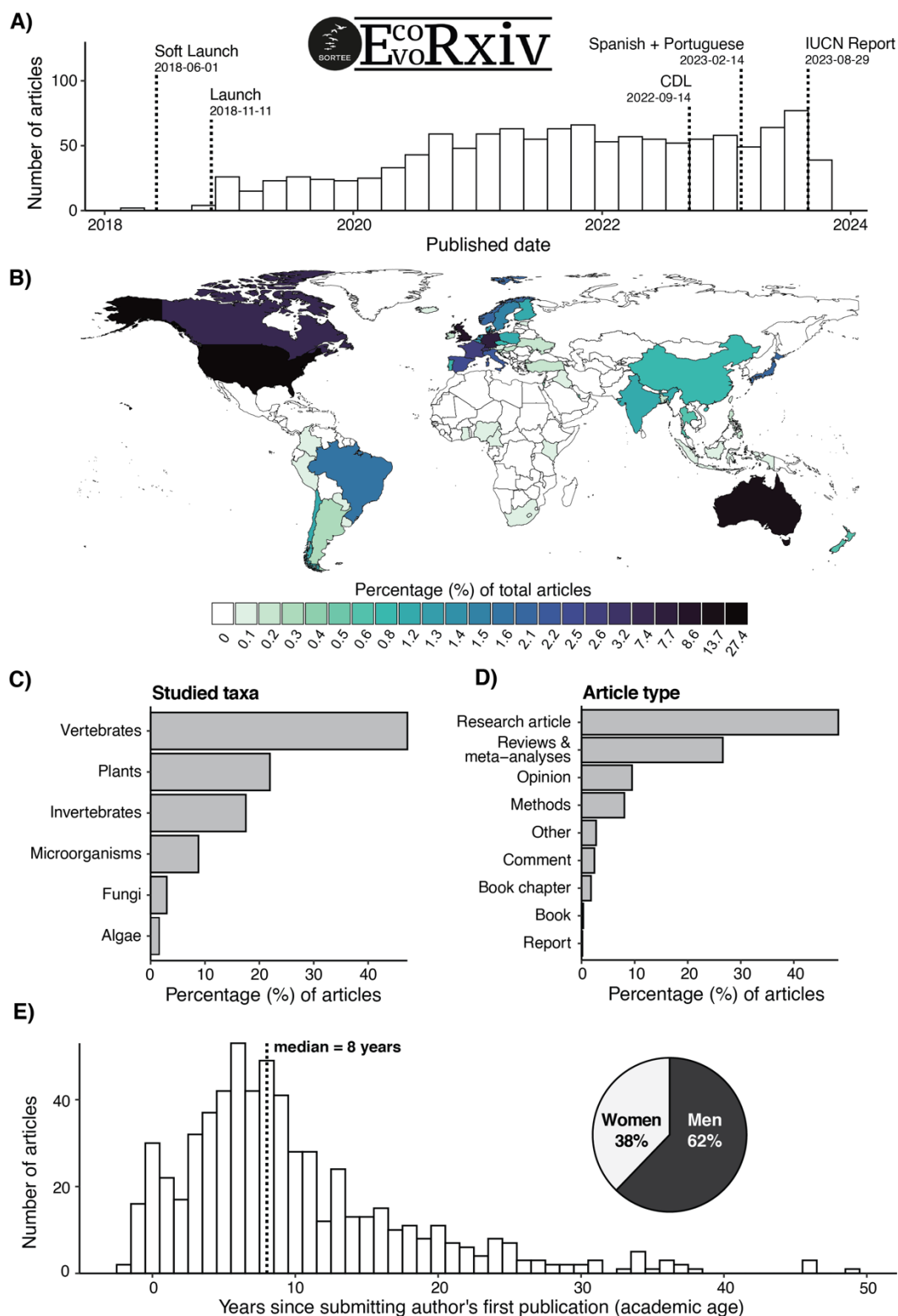


Figure 1- Summary of articles posted to *EcoEvoRxiv*. A) Number of articles (preprints and postprints) published on *EcoEvoRxiv* between 2018 and 2023. *EcoEvoRxiv* was established in June 2018 before the launch in November 2018. Notable milestones include *EcoEvoRxiv* transitioning to the California 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.naturalearthdata.com/about/terms-of-use/>.

## 141 ***Overview of EcoEvoRxiv preprints (and postprints)***

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

## 152 ***Diversifying article types on EcoEvoRxiv: overcoming the ‘grey literature’ problem***

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

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

## 169 ***Breaking down language barriers to scientific communication: improving diversity and*** 170 ***data representation globally***

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

177 *EcoEvoRxiv* is the only preprint server to date that breaks down language barriers to scientific  
178 communication by accepting English, Spanish, Portuguese and French articles. *EcoEvoRxiv* plans to expand

179 to other languages as new non-English editors become available. Such initiatives are incredibly important if  
180 we are to begin filling global voids of scientific knowledge. However, multilingual initiatives have been slow  
181 to take off on *EcoEvoRxiv*, with only a few Spanish articles, and a single Portuguese article, posted since  
182 starting to accept non-English articles in 2023. Part of the challenge in getting authors to submit non-English  
183 articles is the lack of awareness of *EcoEvoRxiv* in non-English speaking countries, cultural differences in the  
184 perception of preprints, and a strong reliance on traditional publishing models that typically mandate  
185 publishing in English<sup>24</sup>.

### 186 ***Generational and gender-based gaps in preprinting practices***

187 Research papers can take a while to be published (see below). However, Early and Mid-Career Researchers  
188 (EMCRs) (~10 years post-PhD) are under pressure to publish rapidly to be competitive in job applications,  
189 promotions, and obtaining grants to progress their careers<sup>7,25</sup>. Preprints are one way EMCRs can achieve  
190 faster dissemination and greater visibility<sup>4</sup>. As such, EMCRs may be expected to make use of preprints more  
191 than colleagues at later career stages. We collected data on the ‘academic age’ of submitting authors by  
192 looking at Google Scholar profiles of authors (when available) and recording their first year of publication in  
193 a peer-reviewed journal. While this is a rough estimate of career stage, there was evidence that the number of  
194 preprints posted decreases with later career stages (negative binomial glm: year slope = -0.1, SE: 0,  $p <$   
195 0.001,  $n = 42$  years). Most preprints were submitted by authors who published their first paper in the last ~10  
196 years (Figure 1E), with the median year since first publication being 2013 (mean = 2010.6; SD = 9.9,  $n =$   
197 1214). These patterns support the expectation that EMCRs may use preprints to make their work more visible  
198 and disseminate their findings more quickly. However, we acknowledge that validating this conclusion does  
199 require more rigorous experimental approaches.

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

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

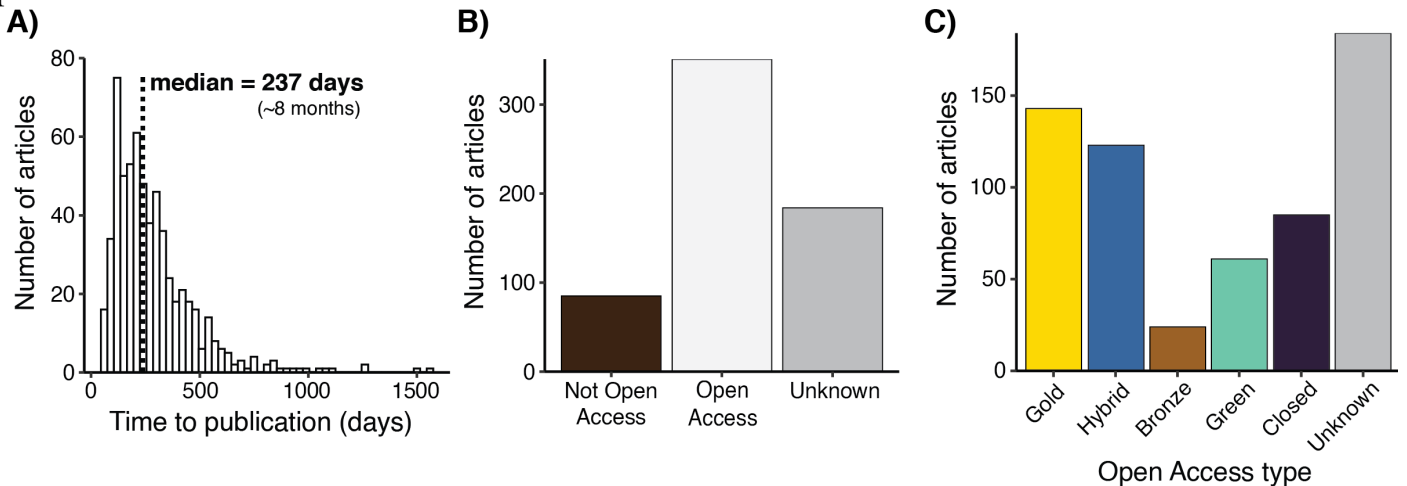


Figure 2- Summary of the publication status of articles on *EcoEvoRxiv*. A) Time between uploading an preprint to *EcoEvoRxiv* and its publication as a peer-reviewed journal article. Articles that were published within 2 months (60 days) of being submitted are considered postprints and are not included 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->.

### *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<sup>7</sup>. 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<sup>7</sup>. Preprints have been proposed as a way to disseminate research more quickly as it can take a long time before results are ultimately published<sup>6,7</sup>. 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 = 1,216) 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 1,549 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.

### *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<sup>6,7</sup>. 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<sup>29</sup>. Most of the published articles were

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

241 Data and code sharing are also key components of open science<sup>30</sup>. In the spirit of ‘openness’, we expected  
242 data and code sharing among preprints to be greater than in many papers published in research journals<sup>30,31</sup>.  
243 Despite this, we found that 54.1% ( $n = 232$ ) of articles relying on data on *EcoEvoRxiv* did not share data, and  
244 56.9% ( $n = 246$ ) did not share code (counting only data-based articles, i.e., excluding reviews, commentaries  
245 or theoretical works).

246 Authors may be reluctant to share data and code for preprints because of the perceived concern that others  
247 may acquire and use their data and code before publication in a journal. Authors of 28.7% ( $n = 123$ ) of  
248 articles that did not share data at the preprint stage did ultimately share data when the article was published;  
249 whereas authors of 35.2% ( $n = 151$ ) never shared data. However, 36.1% ( $n = 155$ ) shared data at both stages.  
250 The same was true for code. Overall, 16.8% ( $n = 72$ ) preprints had no open code at the preprint stage but did  
251 at the published article stage and authors of 45.2% ( $n = 194$ ) preprints did not share code at either stage.  
252 However, 38% ( $n = 163$ ) shared code at both stages. Relatively low code and data-sharing practices in our  
253 sample is consistent with analyses of sharing practices for published articles (e.g.,<sup>31</sup>), even for journals with  
254 strict public data archiving policies<sup>30</sup>.

## 255 **Paving our future to open, transparent and community-driven science**

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

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



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

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

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

## 314 **Acknowledgements**

315 We would like to thank the California Digital Library (CDL) and the CDL team (particularly Alaina  
316 Wrigley) for their support in hosting and maintaining *EcoEvoRxiv* for the Society for Open, Reliable, and  
317 Transparent Ecology and Evolutionary Biology (SORTEE). We would like to thank Gabriela Hidalgo and  
318 Daisy Larios for helping connect us with the IUCN and facilitating discussions to make *EcoEvoRxiv* a place  
319 where IUCN reports can be posted. Finally, we would also like to thank the endless number of SORTEE  
320 volunteers, and those especially on the *EcoEvoRxiv* Committee, who have helped to make *EcoEvoRxiv* a  
321 success. This paper emerged from a hackathon at the 2023 SORTEE conference, and we thank delegates who  
322 attended the session but could not be part of this paper. DWAN would also like to thank the Australian  
323 Research Council for a Future Fellowship (FT220100276). SN and ML are supported by the Australian  
324 Research Council (ARC) Discovery Project Grants (DP210100812 and DP230101248).

## 325 **Data and Code Availability**

326 All data and code can be found on GitHub at: [https://github.com/daniellnoble/ecoevo\\_1000](https://github.com/daniellnoble/ecoevo_1000)

## 327 **Conflict of Interest**

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

## 331 **References**

- 332 1. Ginsparg, P. ArXiv at 20. *Nature* **476**, 145–147 (2011).
- 333 2. Berg, J. M. *et al.* Preprints for the life sciences. *Science* **352**, 899–901 (2016).
- 334 3. Colavizza, G. *et al.* An analysis of the effects of sharing research data, code, and preprints on  
335 citations. *arXiv preprint arXiv:2404.16171* (2024).
- 336 4. Fu, D. Y. & Hughey, J. J. Releasing a preprint is associated with more attention and citations for the  
337 peer-reviewed article. *Elife* **8**, e52646 (2019).
- 338 5. Desjardins-Proulx, P., White, E. P., Adamson, J. J., Ram, K. & Gravel, T. P. D. The case for open  
339 preprints in biology. *PLOS Biology* **11**, e1001563 (2013).
- 340 6. Bourne, P. E., Polka, J. K., Vale, R. D. & Kiley, R. Ten simple rules to consider regarding preprint  
341 submission. *PLOS Computational Biology* **13**, e1005473 (2017).
- 342 7. Vale, R. D. Accelerating scientific publication in biology. *Proceedings of the National Academy of*  
343 *Sciences* **112**, 13439–13446 (2015).
- 344 8. Ni, R. & Waltman, L. To preprint or not to preprint: A global researcher survey. *Journal of the*  
345 *Association for Information Science and Technology* (2024).
- 346 9. Klebel, T. *et al.* Peer review and preprint policies are unclear at most major journals. *PLOS ONE* **15**,  
347 e0239518 (2020).
- 348 10. Chiarelli, A., Johnson, R., Pinfield, S. & Richens, E. Preprints and scholarly communication: An  
349 exploratory qualitative study of adoption, practices, drivers and barriers. *F1000 Research* **8**, 971,  
350 <https://doi.org/10.12688/f1000research.19619.2> (2019).
- 351 11. Fraser, N., Mayr, P. & Peters, I. Motivations, concerns and selection biases when posting preprints: A  
352 survey of bioRxiv authors. *PLOS ONE* **17**, e0274441 (2022).
- 353 12. Guillemaud, T., Facon, B. & Bourguet, D. Peer community in: A free process for the  
354 recommendation of unpublished scientific papers based on peer review. *ELectronic PUBlishing* (2019).
- 355 13. O’Dea, R. E. *et al.* Towards open, reliable, and transparent ecology and evolutionary biology. *BMC*  
356 *biology* **19**, 68 (2021).
- 357 14. Haddaway, N. R. & Bayliss, H. R. Shades of grey: Two forms of grey literature important for reviews  
358 in conservation. *Biological Conservation* **191**, 827–829 (2015).
- 359 15. Haddaway, N. R. *et al.* Eight problems with literature reviews and how to fix them. *Nature Ecology*  
360 *& Evolution* **4**, 1582–1589 (2020).
- 361 16. Amano, T., González-Varo, J. P. & Sutherland, W. J. Languages are still a major barrier to global  
362 science. *PLOS Biology* **14**, e2000933 (2016).

- 363 17. Amano, T. & Sutherland, W. J. Four barriers to the global understanding of biodiversity conservation:  
364 Wealth, language, geographical location and security. *Proceedings of the Royal Society B: Biological*  
365 *Sciences* **280**, 20122649 (2013).
- 366 18. Amano, T., Rios Rojas, C., Boum II, Y., Calvo, M. & Misra, B. B. Ten tips for overcoming language  
367 barriers in science. *Nature Human Behaviour* **5**, 1119–1122 (2021).
- 368 19. Amano, T. *et al.* The role of non-english-language science in informing national biodiversity  
369 assessments. *Nature Sustainability* **6**, 845–854 (2023).
- 370 20. Zenni, R. D. *et al.* Multi-lingual literature searches are needed to unveil global knowledge. *Journal of*  
371 *Applied Ecology* (2023).
- 372 21. Hannah, K., Haddaway, N. R., Fuller, R. A. & Amano, T. Language inclusion in ecological  
373 systematic reviews and maps: Barriers and perspectives. *Research Synthesis Methods* (2024).
- 374 22. White, C. R. *et al.* Geographical bias in physiological data limits predictions of global change  
375 impacts. *Functional Ecology* **35**, 1572–1578 (2021).
- 376 23. Konno, K. *et al.* Ignoring non-english-language studies may bias ecological meta-analyses. *Ecology*  
377 *and Evolution* **10**, 6373–6384 (2020).
- 378 24. Arenas-Castro, H. *et al.* Academic publishing requires linguistically inclusive policies. *Proceedings*  
379 *of the Royal Society B* **291**, 20232840 (2024).
- 380 25. Sarabipour, S. *et al.* On the value of preprints: An early career researcher perspective. *PLOS Biology*  
381 **17**, e3000151 (2019).
- 382 26. Fox, C. W. & Paine, C. T. Gender differences in peer review outcomes and manuscript impact at six  
383 journals of ecology and evolution. *Ecology and Evolution* **9**, 3599–3619 (2019).
- 384 27. Wehner, M. R., Li, Y. & Nead, K. T. Comparison of the proportions of female and male  
385 corresponding authors in preprint research repositories before and during the COVID-19 pandemic. *JAMA*  
386 *network open* **3**, e2020335–e2020335 (2020).
- 387 28. Mullen, L. Predict gender from names using historical data. R package version 0.6.0. (2021).
- 388 29. Jahn, N. Roadoi: Find free versions of scholarly publications via unpaywall. R package version 0.7.2.  
389 <https://github.com/ropensci/roadoi/>, <https://docs.ropensci.org/roadoi/>. (2024).
- 390 30. Roche, D. G., Kruuk, L. E., Lanfear, R. & Binning, S. A. Public data archiving in ecology and  
391 evolution: How well are we doing? *PLOS Biology* **13**, e1002295 (2015).
- 392 31. O’Dea, R. E. *et al.* Preferred reporting items for systematic reviews and meta-analyses in ecology and  
393 evolutionary biology: A PRISMA extension. *Biological Reviews* **96**, 1695–1722 (2021).
- 394 32. Soeharjono, S. & Roche, D. G. Reported individual costs and benefits of sharing open data among  
395 canadian academic faculty in ecology and evolution. *BioScience* **71**, 750–756 (2021).
- 396 33. Roche, D. G. *et al.* Troubleshooting public data archiving: Suggestions to increase participation.  
397 *PLoS biology* **12**, e1001779 (2014).
- 398 34. Anderson, K. R. bioRxiv: Trends and analysis of five years of preprints. *Learned Publishing* **33**, 104–  
399 109 (2020).

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