

## SUPPORTING INFORMATION

Is there really such a thing as *Tropical Biology*?

Emilio M. Bruna<sup>1,2 \*</sup>

<sup>1</sup> Department of Wildlife Ecology and Conservation, University of Florida, PO Box 110430, Gainesville, FL 32611-0430, USA

<sup>2</sup> Center for Latin American Studies, University of Florida, PO Box 115530, Gainesville, FL 32611-5530, USA

\* Corresponding author; email: embruna@ufl.edu.

<sup>1</sup> **1. Collection, processing, and visualization of bibliometric data**

<sup>2</sup> To identify the conceptual domains studied by researchers working in ‘Tropical’ and  
<sup>3</sup> “non-Tropical” locations, I used information extracted from the bibliographic records of  
<sup>4</sup> articles published These studies were published from 1990-2022 in N = 10 journals  
<sup>5</sup> (*Journal of Evolutionary Biology*, *Ecology*, *Journal of Applied Ecology*, *Evolution*,  
<sup>6</sup> *Biotropica*, *Journal of Ecology*, *Tropical Conservation Science*, *American Naturalist*,  
<sup>7</sup> *Tropical Ecology*, *Journal of Tropical Ecology*). Specifically, I compared (1) article  
<sup>8</sup> keywords, (2) individual words in article titles (e.g., *seed*, *species*), and (3) title bigrams  
<sup>9</sup> (i.e., pairs of sequential words in titles, e.g., *seed predation*, *species diversity*). Below I  
<sup>10</sup> describe how the article records were identified, downloaded, processed, and assigned to the  
<sup>11</sup> ‘Tropical’ and “non-Tropical” categories using code written in the R programming language  
<sup>12</sup> (R Core Team 2023).

<sup>13</sup> On 8 February 2023, I downloaded all bibliographic data available in SCOPUS and  
<sup>14</sup> the Web of Science ‘Core Collection’ for all articles published in the focal journals; both  
<sup>15</sup> SCOPUS and the Web of Science were queried because they differ in the years indexed for  
<sup>16</sup> each journal. I then used the `refsplitr` package (Fournier *et al.* 2020) to process the  
<sup>17</sup> records and remove any duplicates. After removing all stopwords (Benoit *et al.* 2021) from  
<sup>18</sup> article titles and keywords, I spell-checked, stemmed, and lemmatized all of the keywords  
<sup>19</sup> and title words. I also extracted bigrams from titles with the `tidytext` library (Silge &  
<sup>20</sup> Robinson 2016). Finally, I identified each article as either ‘Tropical’ or ‘non-Tropical’; all  
<sup>21</sup> articles published in (*Journal of Evolutionary Biology*, *Ecology*, *Journal of Applied Ecology*,  
<sup>22</sup> *Evolution*, *Biotropica*, *Journal of Ecology*, *Tropical Conservation Science*, *American*  
<sup>23</sup> *Naturalist*, *Tropical Ecology*, *Journal of Tropical Ecology*) were assigned to the ‘Tropical’  
<sup>24</sup> category, while articles published in the other journals were assigned to one of these  
<sup>25</sup> categories based on a search of the titles, keywords, or abstracts for a list of  
<sup>26</sup> domain-specific terms (e.g., tropical: *amazon*, *andes*, *congo*, *bci*, *chamela*; non-tropical:  
<sup>27</sup> *finland*, *boreal*, *eastern decid*, *arctic*, *polar*). These procedures resulted in N = 37,807 total

28 articles published, of which N = 11,210 reported research conducted in the tropics and N =  
29 26,597 were based on work conducted in other locations. Collectively, these articles used N  
30 = 62,883, N = 25,207 unique title words, and N = 126,796 title bigrams.

31 The number of articles varies widely between journals, as does the number of  
32 keywords per article. Comparing counts of keyword frequency in tropical and non-tropical  
33 articles could therefore bias results towards the content published a small number of  
34 journals. To correct for this, I calculated the percentage of articles in each geographic  
35 category that using each keyword, title word, or bigram. I then selected the N = 50 most  
36 frequently used terms in each geographic category, and identified (a) any terms that  
37 ‘tropical’ and ‘non-tropical’ articles had in common, and (b) any terms that were unique to  
38 each article category.

## 39 **2. Data and Code**

40 The version of the code used to review, correct, and prepare the data set (version 1.0.0) is  
41 available at Zenodo at <*DOI added upon acceptance*>, and the data set used in this  
42 publication is available in Dryad at <*DOI added upon acceptance*>. The code used to  
43 prepare this publication, including statistical summaries reported in the text, tables, and  
44 figures, is also available at Zenodo at <*DOI added upon acceptance*>. Questions regarding  
45 the data or code, or suggestions for improvement should be posted as Issues on the  
46 project’s Github Repository ([https://github.com/BrunaLab/atbc2022\\_plenary\\_talk](https://github.com/BrunaLab/atbc2022_plenary_talk)) or  
47 referred to E. M. Bruna. Summaries of any post-publication updates will be posted to the  
48 NEWS.md file of this Github Repository.

## 49 **REFERENCES**

- 50 BENOIT, K., D. MUHR, and K. WATANABE. 2021. Stopwords: Multilingual  
51 stopword lists. <https://CRAN.R-project.org/package=stopwords>
- 52 FOURNIER, A. M. V., M. E. BOONE, F. R. STEVENS, and E. M. BRUNA. 2020.  
53 Refsplitr: Author name disambiguation, author georeferencing, and mapping of

- 54 coauthorship networks with Web of Science data. Journal of Open Source Software 5: 2028.
- 55 R CORE TEAM. 2023. R: {A} language and environment for statistical computing.
- 56 R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- 57 SILGE, J., and D. ROBINSON. 2016. Tidytext: Text mining and analysis using tidy
- 58 data principles in R. Journal of Open Source Software 1(3).

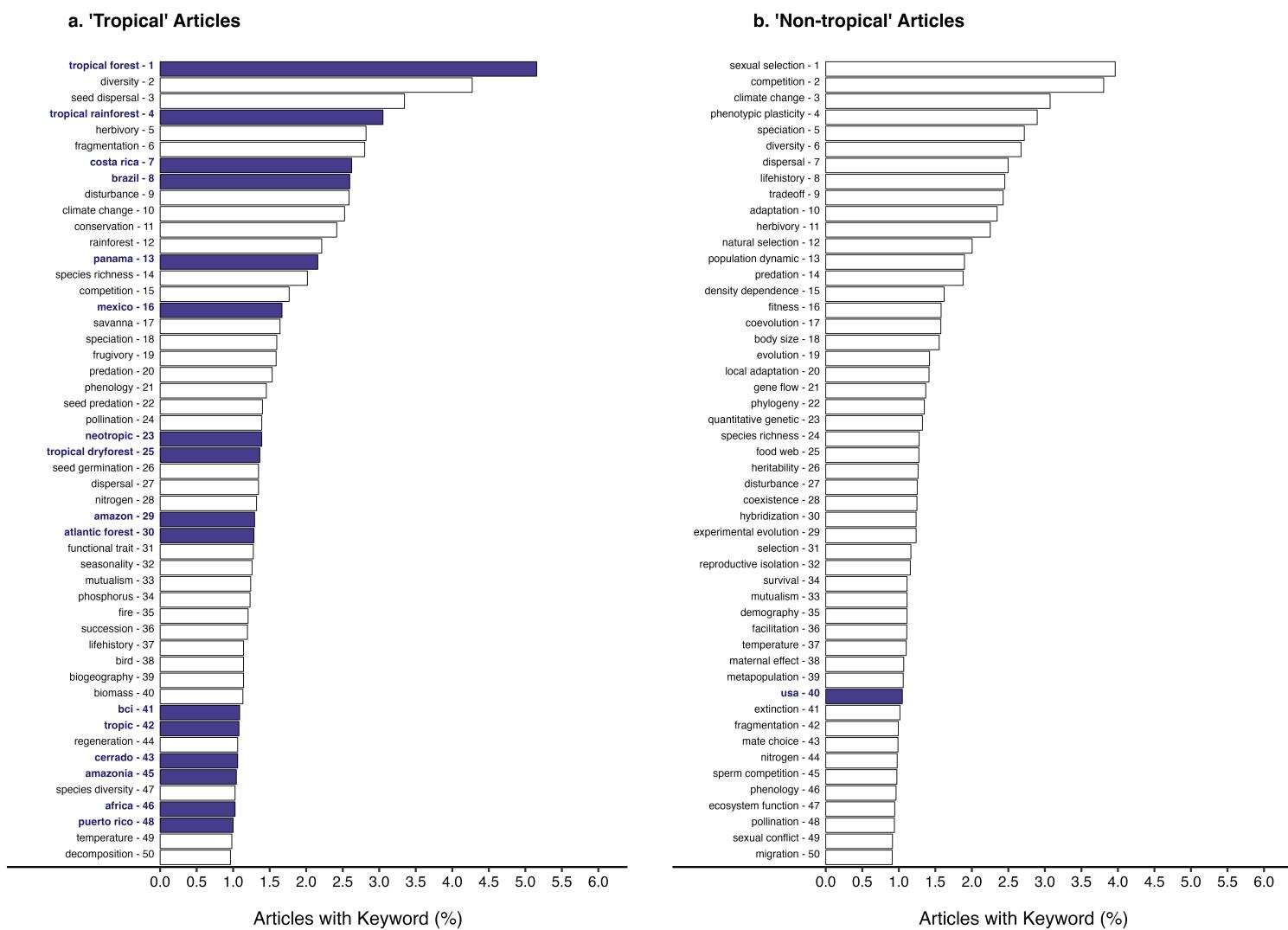


Figure S1. The N = 'r cutoff' most common keywords from articles based on research conducted in (a) the tropics and (b) non-tropical regions. The rank of these words is based on the percentage of articles in each category that included them. Terms reflecting geography (e.g., *tropics*, *Peru*, *Southern*) are indicated in bold and with filled bars.

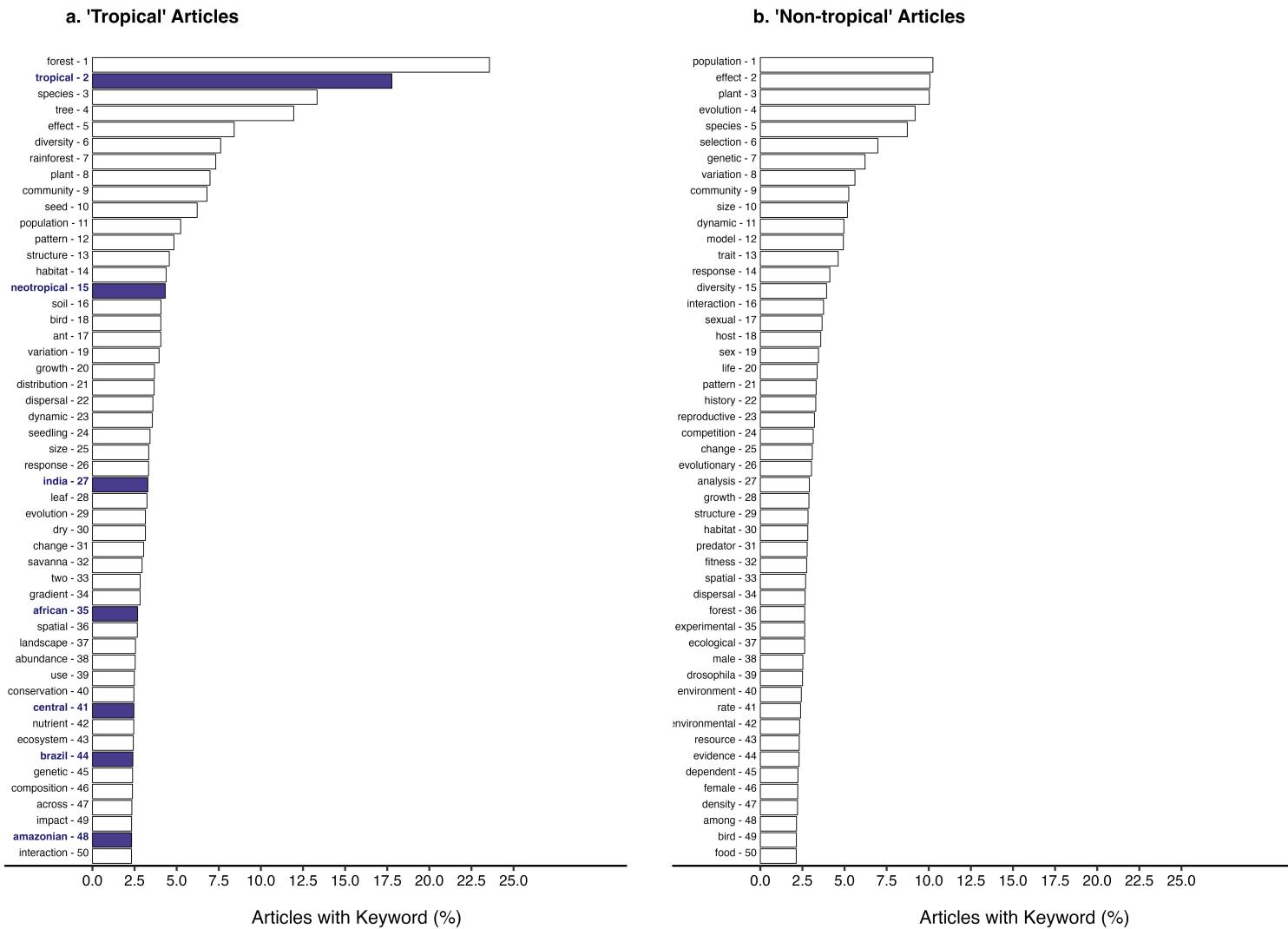


Figure S2. The N = 'r cutoff' most common words in the titles of articles based on research conducted in (a) the tropics and (b) non-tropical regions. The rank of these words is based on the percentage of article titles in each category that included those words. Terms reflecting geography (e.g., *tropics*, *Peru*, *Southern*) are indicated in bold and with filled bars.

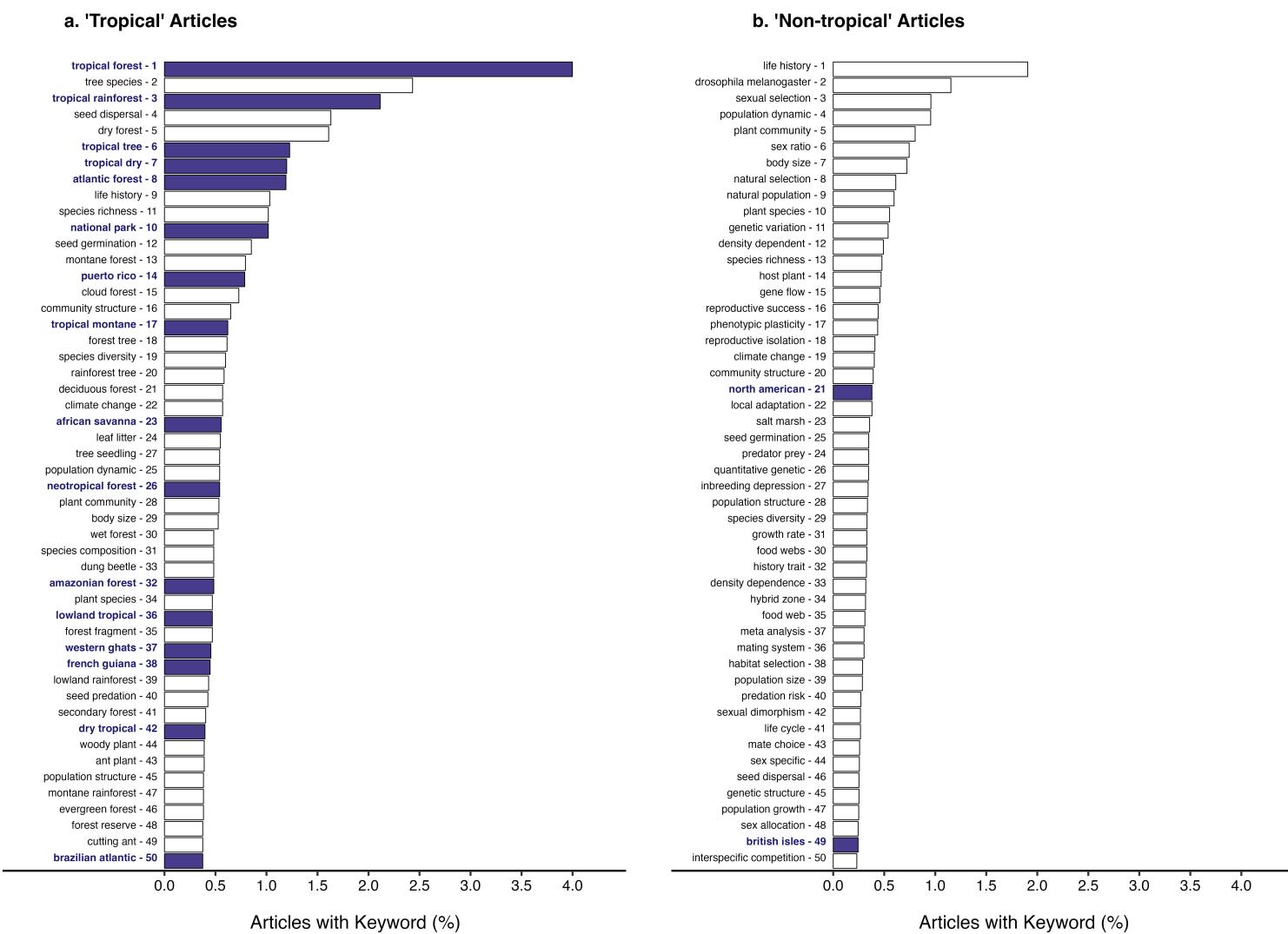


Figure S3. The N = ‘r cutoff’ most common bigrams in titles of articles based on research conducted in (a) the tropics and (b) non-tropical regions. The rank of these words is based on the percentage of article titles in each category that included those words. Bigrams reflecting geography (e.g., *tropics*, *Peru*, *Atlantic Forest*) are indicated in bold and with filled bars.

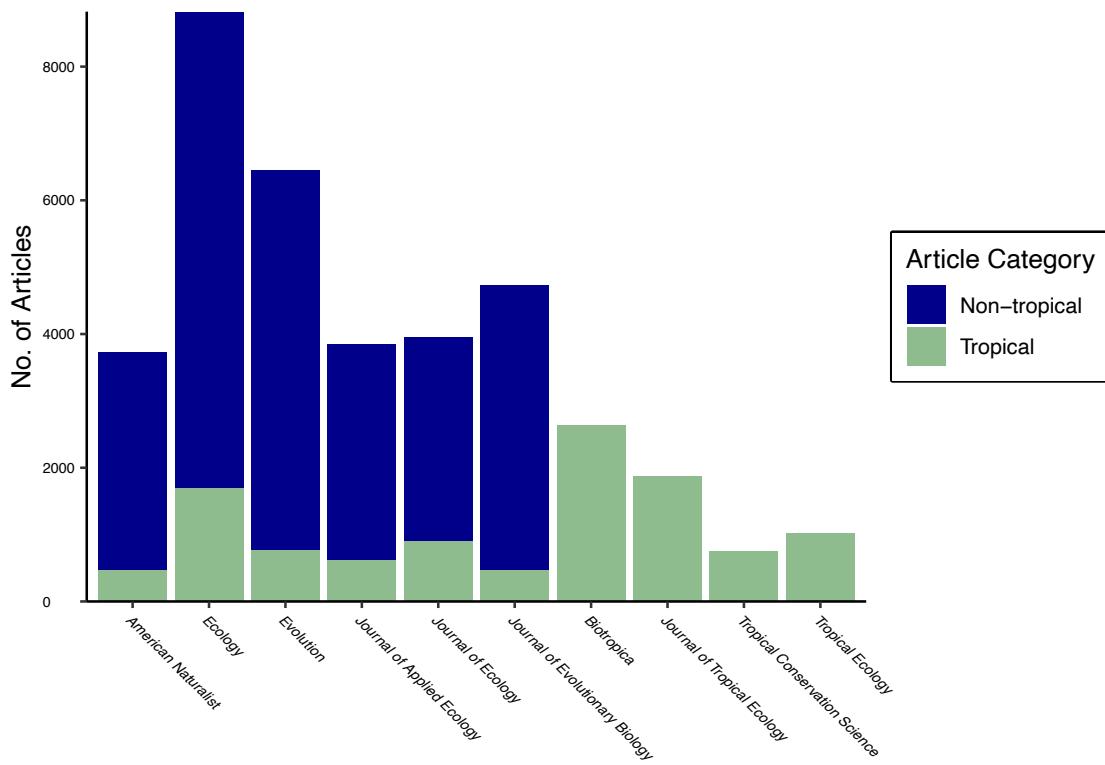


Figure S4. The number of articles from each journal and geographic category that were used in used the analysis of keywords.

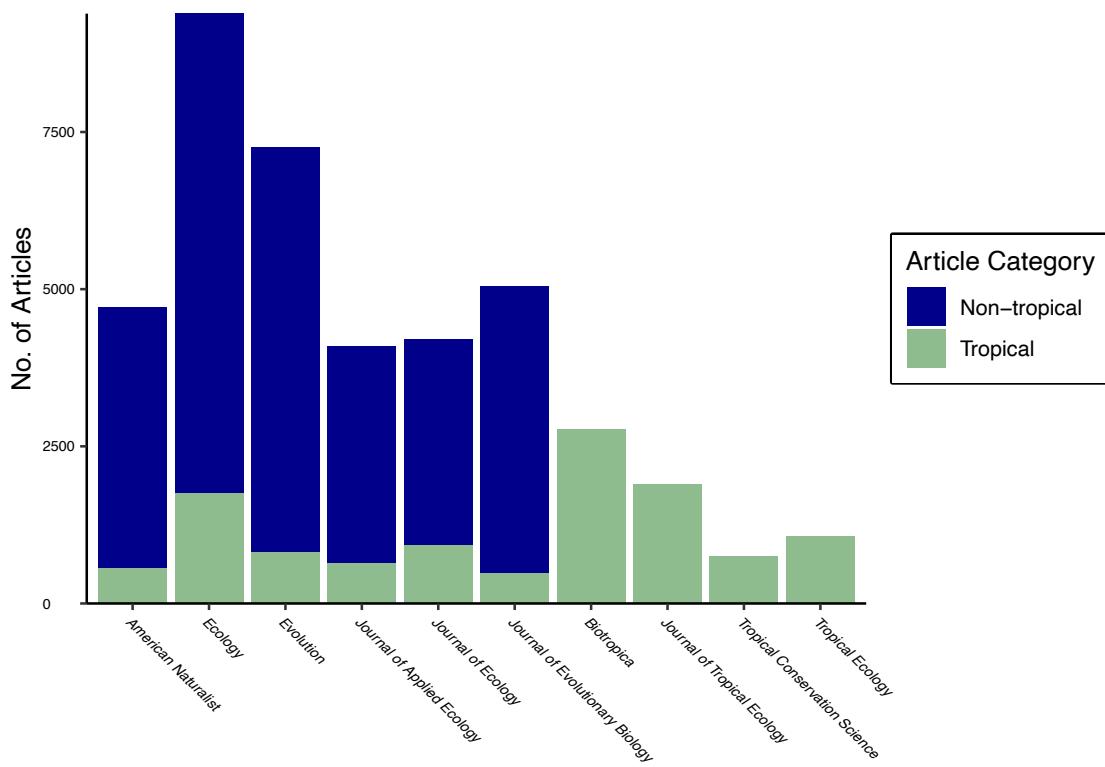


Figure S5. The number of articles from each journal and geographic category that were used in the analysis of title words and title bigrams.