Gendered male and high-income country authors dominate publication at a One Health research organization

Cecilia A. Sánchez[1,#a*], Collin J. Schwantes[1,#a], Shannon Ball[1], Siyeun Kim[1], Sarah Munro[1,#b], Rebecca Leaman[1,#c], Cadhla Firth[1,#d]

[1]EcoHealth Alliance, New York, NY, United States of America

[#a]Current address: Department of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, CT, United States of America

[#b]Current address: Department of Biological Sciences, University of Arkansas, Fayetteville, AR, United States of America

[#c]Current address: Columbia Global, Columbia University, New York, NY, United States of America

[#d]Current address: College of Public Health, Medical and Veterinary Sciences, James Cook University, Townsville, Queensland, Australia

[*]Corresponding author

Email: cecilia.sanchez@yale.edu (CAS)

Abstract

Authorship on academic publications carries substantial weight for researchers in science fields. One's position in a list of authors is typically used to signal information about author contributions and status, with the first and last authorship positions regarded as the most prestigious and important for career advancement. Therefore, any inequities that exist in the allocation of publication authorship (e.g. due to gender or geography) could affect researchers' career progression. We assessed patterns in publication authorship at EcoHealth Alliance, a non-profit organization that conducts One Health and conservation research. We compiled a corpus of 451 peer-reviewed journal articles published from 2011-2022, each of which had at least one EcoHealth Alliance-affiliated author, and gathered information on the gender and country affiliation of first and last authorships (FLAs). We found that gendered male authors represented ~60% of authors, 65% of FLAs, and 91% of highly productive authors. Last authorships were particularly male-dominated, with ~2.7 times as many last authorships by gendered male authors as by gendered female authors. Gendered male authors were more structurally important to the author network on average and comprised 65% of highly "powerful" authors in the network. High-income countries were also overrepresented in the corpus, with ~72% of FLAs listing a high-income country affiliation. We conclude by offering recommendations for researchers, organizations, and funders and publishers to improve equity in authorship practices.

Introduction

Authorship on academic publications carries substantial weight for researchers in science, technology, engineering, and mathematics (STEM), as it can bestow prestige, bolster reputations, and influence career trajectories (1). An individual's publication record is a key consideration in hiring, promotion, and tenure decisions (2-4). Greater productivity in the form of authorship has also been linked to obtaining more research funding (5, 6) and higher pay for STEM researchers (7). However, the benefits of publication are not experienced equitably by those who contribute to a manuscript. In the sciences, authorship position (i.e. one's position in a list of authors on a publication) is typically used to signal information about author contributions and status, with the first and last positions regarded as the most prestigious (8, 9) and thus particularly important for career advancement (4, 10). Therefore, any inequities that exist in the allocation of publication authorship separate from actual contribution (e.g. due to gender or geography) could affect researchers' career progression, potentially exacerbating existing inequalities in STEM (11, 12).

Multiple studies have demonstrated that the perceived gender and affiliated country of an author impacts publication authorship. Though differences exist between STEM fields, women are generally under-represented as authors, in overall number of authorships, and in first and last authorships (13-15). For instance, a global study of scientific output found that women accounted for less than 30% of authorships and that for every article with a female first author, there were nearly two by a male first author (13). Although gendered authorship gaps have improved in recent decades, change is slow and disparities persist (14-16). There is a widening gender gap in the last authorship position (14), which appears to have been exacerbated by the COVID-19 pandemic (17). Though it is difficult to determine the mechanisms for gendered authorship disparities, suggested explanations include the slower career progression and/or

attrition of women in science, an influence of gender in authorship negotiations, and differences in attribution and recognition of women's contributions (14, 16, 18, 19). Notably, authors from low- and middle-income countries are also less likely to be listed as first or last authors on publications compared to authors from high-income countries (20-22). As a result, researchers with multiple marginalized identities are especially underrepresented as authors and in prestigious authorships. In a study of all publications by The Nature Conservancy (TNC), one of the world's largest conservation non-profit organizations, women in the Global South represented just 3% of all authors while first and last authorships by women in the Global South each comprised less than 1% of all authorships (15).

Research organizations that work across geographies and STEM fields provide a unique opportunity to examine the influence of gender and geography on publication authorship. EcoHealth Alliance (EHA) is a non-profit organization that, according to its mission, "develops science-based solutions to prevent pandemics and promote conservation" (https://www.ecohealthalliance.org/). Founded in 1971 as Wildlife Trust, the organization rebranded in 2010 to reflect a transition from a conservation focus to a broader "One Health" agenda, promoting the health of humans, animals, and the environment to protect the public from emerging infectious diseases. EHA is headquartered in New York City, but partners with universities, non-profit organizations, and local governments, typically in low- and middle-income countries. At the time this project was initiated (mid-2023), EHA employed approximately 50 scientific staff with a range of expertise and training (e.g. ecologists, biologists, veterinarians, public health professionals, policy experts) and 10 administrative staff. While employed at EHA, several authors of this manuscript formed and participated in a Women in Science staff affinity group. The goals of the group were to provide a sense of belonging and

community for female staff, offer opportunities for networking, mentoring, and professional development, and ultimately improve recruitment, advancement, and retention of women at EHA. Coincident with discussions within the group and EHA about diversity, equity, and inclusion, we initiated a project to examine the role of gender and geography in scientific publications authored by EHA-affiliated scientists. We were particularly motivated by the case study of TNC publications, which found that only 36% of authors on TNC publications were women, and that women made up just 31% and 24% of first and last authorships, respectively (15). Additionally, 87% of authors in their dataset were located in the Global North (Western Europe, Northern America, Australia, New Zealand, and Israel). We aimed to similarly assess past publishing practices within EHA, which could then be used as a baseline against which to measure future change.

Materials and methods

Ethical and organizational approval

On June 17, 2023, this study was approved as exempt by the Health Media Labs Institutional Review Board (Protocol #2264). This study was also approved by EHA senior leadership. Data were accessed for research purposes from June 18, 2023 through June 4, 2024. Authors had access to information that could identify individual participants during and after data collection.

Data collation and cleaning

Prior to initiating this study, EHA developed a catalog of research outputs associated with the organization for use in tracking internal metrics. Research outputs consisted mainly of scientific journal articles, but also included materials such as conference abstracts, book chapters, datasets, reports, and preprints. For a research output to be included in the catalog, at least one of its authors must have listed an affiliation with EHA or Wildlife Trust (the organization's former name). Research outputs and associated authorship records were imported into this catalog via an OpenAlex API query using the *openalex* R package (23, 24), when the authorship institution attribute for at least one author contained the Research Organization Registry (https://ror.org/) identifier for EHA. Records were processed digitally to keep pertinent fields, identify potential duplicates, and store them in an Airtable (https://airtable.com/) database. Each research output is associated with metadata such as title, publication date, author names, and author affiliations (which typically include an organization and country). This catalog is maintained on the Airtable platform and is accessible to EHA employees.

For this study, we exported a copy of EHA's research outputs catalog to a separate Airtable database, accessible only to project personnel. We filtered the dataset to include only peer-reviewed journal articles published from January 1, 2011 to December 31, 2022. We chose the start date to align with the shift in the organization's name and research focus. In some cases, it was unclear whether an article was peer-reviewed (e.g. opinion pieces are peer-reviewed in some journals but not others); we referred to journal policies and excluded articles if we were unable to verify that peer review had occurred. We note that "gray" and "white" literature (e.g. graduate theses, government reports, policy documents, technical reports) also represent important research outputs and require similar levels of time and effort to produce as peer-

reviewed research. However, we chose to focus on peer-reviewed literature following past work (15) and because gray and white literature may have their own, different authorship norms. Records were manually reviewed and cleaned if necessary to ensure accuracy. Henceforth, we refer to this set of peer-reviewed, EHA-affiliated journal articles as the *corpus*.

From the corpus, we developed three related tables: an articles table, an authors table, and an authorships table. Here, we use the term *author* to refer to a unique individual, and *authorship* to refer to an individual's contribution to a specific journal *article*. We cleaned the authors table by removing duplicates, referring to sources such as ORCID (Open Researcher and Contributor ID), ResearchGate, and Google Scholar profiles to ensure that unique authors were identified as such despite minor differences in name spelling, use of initials, or accents. We also collected information on author gender (see *Gender classification of authors* below). We classified authorships as either first (including sole authorships), middle, or last, and restricted the authorships table to include only first and last authorships in recognition of the greater contribution and prestige associated with these positions (9). In the case of co- first authorships, we included only the author listed first.

We noted the country listed for each authorship affiliation (hereafter, *authorship geography*). Sometimes, a single authorship listed multiple affiliations; in these cases, we selected only the first affiliation, assuming it represented an organization or location of primary importance to the author. We classified each country as low-income, lower-middle-income, upper-middle-income, or high-income using historical World Bank data (https://datacatalogfiles.worldbank.org/ddh-published/0037712/DR0090754/OGHIST.xlsx). Because the World Bank classifies incomes annually, we used the most common designation for a country over the period of our dataset. For example, Bangladesh was classified by the World

Bank as low-income from 2011-2013 and lower-middle-income from 2014-2022; therefore, we classified it as lower-middle-income. Finally, we assessed whether each article had a geographic focus; that is, the research described in the article required the authors' physical presence in a country (e.g. for fieldwork) or specialized knowledge about a country. If yes, we recorded all focal countries of an article as the *article geography*. If the work described in an article could be performed regardless of location (e.g. a review or model-based article), we classified the article geography as "non-specific".

Gender classification of authors

We employed a two-pronged approach to gather author gender information (for authors with at least one first or last authorship; n = 498 authors). We note that gender identity can be a sensitive topic. Individuals may have different gender presentations in their personal and professional lives, and gender identity can change over time. As such, all gender-related data were stored only in our project database, with access restricted to project personnel. We deposited a de-identified database for public use on Zenodo (25).

1) Pronouns-based approach. Data on self-identified genders are typically not publicly available. However, there is a growing practice of sharing one's pronouns (e.g. she/her, he/him, they/them, she/they, they/he) in conversation and online (e.g. professional websites, email signatures) (26, 27). We therefore annotated author data with manually-gathered pronouns (28) to infer gender. To gather author pronouns, we drew from professional interactions we have had with authors and publicly available online information (e.g. lab websites, interviews, press releases, conference programs). Pronouns were recorded as "she/her/hers", "he/him/his", "they/them/their" (with the option to select multiple of these) or "unable to find". We classified

authors as *gendered female* if we only found evidence that they used she/her/hers pronouns, as *gendered male* if we only found evidence that they used he/him/his pronouns, and as *gendered nonbinary* if we found evidence they used they/them/their pronouns or any combination of he/she/they pronouns (29). Here, we use *nonbinary* as "an umbrella term for people whose gender identity doesn't sit comfortably with 'man' or 'woman'" (30). We follow Van Buskirk and colleagues (31) in using the terms *gendered male* and *gendered female* to emphasize that these are externally imposed classifications, and acknowledge that a person's gender does not necessarily correspond to the pronouns they use. However, using pronouns as a proxy for gender identity allowed us to include those with nonbinary identities, who are often excluded in similar analyses (15, 32).

2) Name-based approach. We hypothesized that our ability to find author pronouns would be diminished for authors outside the United States, given that sharing pronouns is not a global practice. Therefore, we also used the *nomquamgender* python package (31) to assign a probability *p(gf)* that each author was gendered female. The package uses a "dictionary" of name-gender associations from more than 150 countries to assign a *p(gf)* value to an individual based on their name. A user can then set a threshold to classify binary gender based on *p(gf)*. We used the default threshold of 0.1, meaning we classified an author as *gendered female* if *p(gf)* was ≥ 0.90 and as *gendered male* if *p(gf)* was ≤ 0.10. If *p(gf)* was between 0.10 and 0.90, we classified an author as *uncertain*. Names that do not occur in the dictionary of name-gender associations cannot be assigned *p(gf)* values; we classified these names as *undetermined*.

We assessed concordance between gender classifications made using the pronouns-based approach versus the name-based approach and made a final gender classification list by merging the results of our two approaches, deferring to gender based on pronouns in cases of

disagreement. When we were unable to determine gender by either approach, we classified author gender as *unknown*. We acknowledge that all methods for classifying author gender are imperfect, and we may have incorrectly classified the gender of some authors. However, by employing our two-pronged approach, we were able to maximize the data available for analysis and include genders beyond the binary.

Analyses of gender, geography, income, and authorship

All analyses were performed in the R statistical environment v 4.4.2 (33). To examine patterns in authorship positions by gender, we calculated the percent of all first and last authorships (FLAs) across all combinations of authorship position and gender. We also calculated the number of FLAs associated with each unique author, and plotted histograms of FLAs by author gender to explore whether author productivity differed by gender. For the most productive authors, we calculated the percent of their FLAs composed of last authorships.

To examine economic patterns in authorship positions, we calculated the percent of all FLAs across all combinations of authorship position and country income. We compared the number of authorships by gendered male and gendered female authors within each country to examine interactions of geography and gender. Further, we calculated the percent of all FLAs across all combinations of authorship position, gender, and country income.

We used a linear model to examine effects of authorship position, country income, and year on the annual percent of FLAs by gendered female authors. Explanatory variables for the model included an interaction between year and authorship position, an interaction between country income and authorship position, and main effects of year, authorship position, and country income. We expected that the percent of authorships by gendered female authors would

increase over time, but the rate of increase would be lower for last authorships. Last authorships are often reserved for senior scientists (e.g. principal investigators), and promotion to this senior role typically takes years. Based on preliminary visual exploration of the data, we also expected the relationship between authorship position and the percent of authorships by gendered female authors to differ for countries of different income. Due to limited authorship data for low-income, lower-middle-income, and upper-middle-income countries in our dataset (see Results), we collapsed these three categories into one and treated country income as a binary variable in the model (high-income versus low- and middle-income). We treated year as a continuous variable, and centered it around 2011 (the first year in the dataset) to improve coefficient interpretability. Authorship position was a binary variable.

For research taking place outside of the United States (where EHA is located), we explored the association between the geographic focus of an article and the geographic affiliations of the first and last authors. We did this to assess how often locally-based researchers received credit in the form of prestigious first and last authorships for their critical roles in these research projects. We first restricted the table of articles to exclude those with a “non-specific” article geography and those with an article geography that included the United States. We then calculated how often authorship geography "matched" article geography for the i) first authorship position only, ii) the last authorship position only, iii) either authorship position, and iv) both authorship positions. If multiple countries were listed for article geography (e.g. because fieldwork took place in several locations), we counted a match if the authorship geography was the same as any of those countries. To understand if authorship practices changed over time, we then repeated these calculations for two time periods: 2011-2016, and 2017-2022.

Network analyses to examine gender and author centrality

Network analyses are commonly used to reveal structural aspects of social relationships, and centrality measures are designed to identify individuals within a network who are important to its structure (34). In the context of co-authorship networks, these individuals tend to be senior researchers or highly-cited individuals (35). To examine collaborations between authors, we calculated centrality measures for the network of all FLAs in the corpus. In this analysis, articles and authors represent two components of a bipartite graph, where authors are linked by co-authorship on an article. Because we were interested in connections between authors, we re-projected the bipartite graph such that it contained weighted edges between authors (nodes) based on their co-authorships. We then used the author network to explore the relationship between gender and two measures of centrality: *betweenness centrality* and *harmonic centrality*.

Betweenness centrality measures the number of shortest paths on which a node resides (36), where the shortest path is the walk between two nodes that requires traversing the least number of edges. A node with high betweenness centrality is structurally important to the graph. In our analysis, authors with high betweenness centrality likely represent people in positions of "power" within the network (e.g. principal investigators, those who control resources). Harmonic centrality measures the degree of a node (i.e. number of connections to other nodes) and its neighbors, up to a certain distance (37). The higher the degree of a node and its neighbors, the higher the harmonic centrality–with the important caveat that the influence of neighbors decays with distance and is not inflated for unconnected subcomponents of the graph (34). This provides information about how connected nodes and their neighbors are to the rest of the graph. In our analysis, an author with high harmonic centrality is collaborating with many people who are also collaborating with many people.

We calculated betweenness centrality and harmonic centrality for each author using the *igraph* R package (38, 39). We also calculated mean betweenness centrality and mean harmonic centrality for each gender group, as well as 95% high density confidence intervals (HDCIs) for each measure. We calculated 95th percentile values for betweenness centrality and harmonic centrality and used these as cutoffs to tally the number of highly powerful and highly collaborative authors according to gender. Finally, we created two depictions of the authorship network using node color to represent author gender and node size to represent each centrality measure.

Results

Dataset summary

Filtering the EHA research outputs catalog to peer-reviewed journal articles published from January 1, 2011 to December 31, 2022 resulted in a corpus of 451 articles. We identified 898 FLAs associated with those articles, which were linked to 498 unique authors.

Comparison of two approaches to classify author gender

Using the pronouns-based approach, we classified 280 authors (56.2%) as gendered male, 181 (36.3%) as gendered female, and 1 (0.2%) as gendered nonbinary. We were unable to find pronouns for 36 authors (7.2%). Of the 462 authors whose pronouns were identified, publicly available online information was the source for 323 (69.9%), while professional interactions were the source for the remaining 139 (30.1%). Using the name-based approach, we classified 265 authors (53.2%) as gendered male, 159 (31.9%) as gendered female, 56 (11.2%) as

uncertain, and 18 (3.6%) as undetermined. Agreement between the two gender classification approaches was generally high (Table S1); out of 498 authors, both approaches classified 151 authors as gendered female and 244 authors as gendered male. There were four cases where an author was gendered female or gendered nonbinary based on pronouns but gendered male based on name, and three cases where an author was gendered male based on pronouns but gendered female based on name.

Gender composition of authors and authorships

After combining the results of our two gender classification approaches, prioritizing gender inferred using pronouns in cases of dataset disagreement, we ultimately classified 297 authors (59.6%) as gendered male, 186 (37.3%) as gendered female, 1 (0.2%) as gendered nonbinary, and 14 (2.8%) as unknown gender. Of all FLAs, 584 (65.0%) were by gendered male authors, 295 (32.9%) were by gendered female authors, 1 (0.1%) was by a gendered nonbinary author, and 18 (2.0%) were by authors of unknown gender. Our gender classification process was least successful for authorships listing an affiliation with China, with 16 out of 44 authorships for this country classified as unknown gender. Given that only one author was classified as gendered nonbinary, we included that author in figures except where noted, but did not draw conclusions about how publication practices at EHA impact nonbinary authors.

For both the first and last authorship positions, there were more authorships by gendered male authors than by gendered female authors (first-male: 29.3% of all FLAs, first-female: 19.5%, last-male: 35.7%, last-female: 13.4%; Fig. S1). Nearly three-quarters of all authors (73.7%) had just one authorship each (226 gendered male authors and 141 gendered female authors) (Fig. 1). There was only one gendered female author with ≥ 10 total FLAs, while there

were ten gendered male authors with 10 - 44 total FLAs each. Collectively, those ten gendered male authors accounted for 187 (20.8%) of all FLAs in the corpus (Fig. 1). Last authorships made up >65% of FLAs for eight of the eleven most productive authors (Table S2).

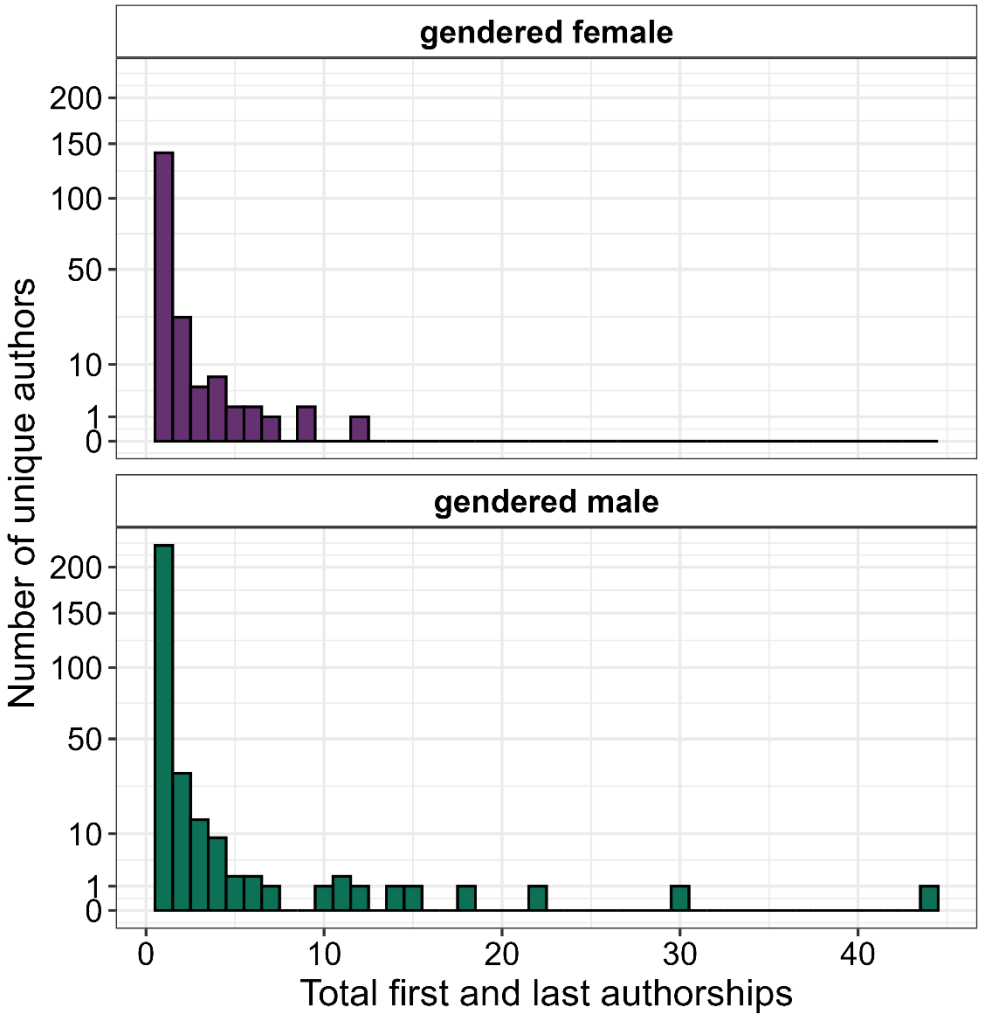


Fig 1. Total first and last authorships associated with unique authors. Gendered female (n = 186) and gendered male (n = 297) authors are displayed in the top and bottom panels respectively. Authors classified as nonbinary (n = 1) or unknown (n = 14) are not displayed. Note that the y-axis is on a square-root scale.

Intersections of gender, country income, and authorship position

A total of 43 countries (high-income: n = 19, upper-middle-income: n = 8, lower-middle-income: n = 13, low-income: n = 3) were represented in authorship affiliations (Fig. S2). Most authorships listed an affiliation with the United States (49.1%), Bangladesh (12.1%), Australia (9.2%), China (4.9%), or the United Kingdom (4.2%), together comprising 79.5% of all FLAs. Within each of these five countries, there were more authorships by gendered male authors than authors of any other gender (Fig. S3). Most FLAs had a high-income country affiliation (71.5%), while other income groups were less well represented (upper-middle: 12.0%, lower-middle: 15.1%, low: 1.3%; Fig. S4).

For authorships with a high-income country affiliation, there was an interplay between gender and authorship position (Fig. 2). Specifically, there were more first authorships than last authorships (15.9% versus 9.4% of all FLAs) by gendered female authors, while there were more last than first authorships (26.7% versus 19.2% of all FLAs) by gendered male authors. The data on authorships with an upper-middle-income, lower-middle-income, or low-income country affiliation were too limited to determine if differences by gender and authorship position were meaningful (Fig. 2).

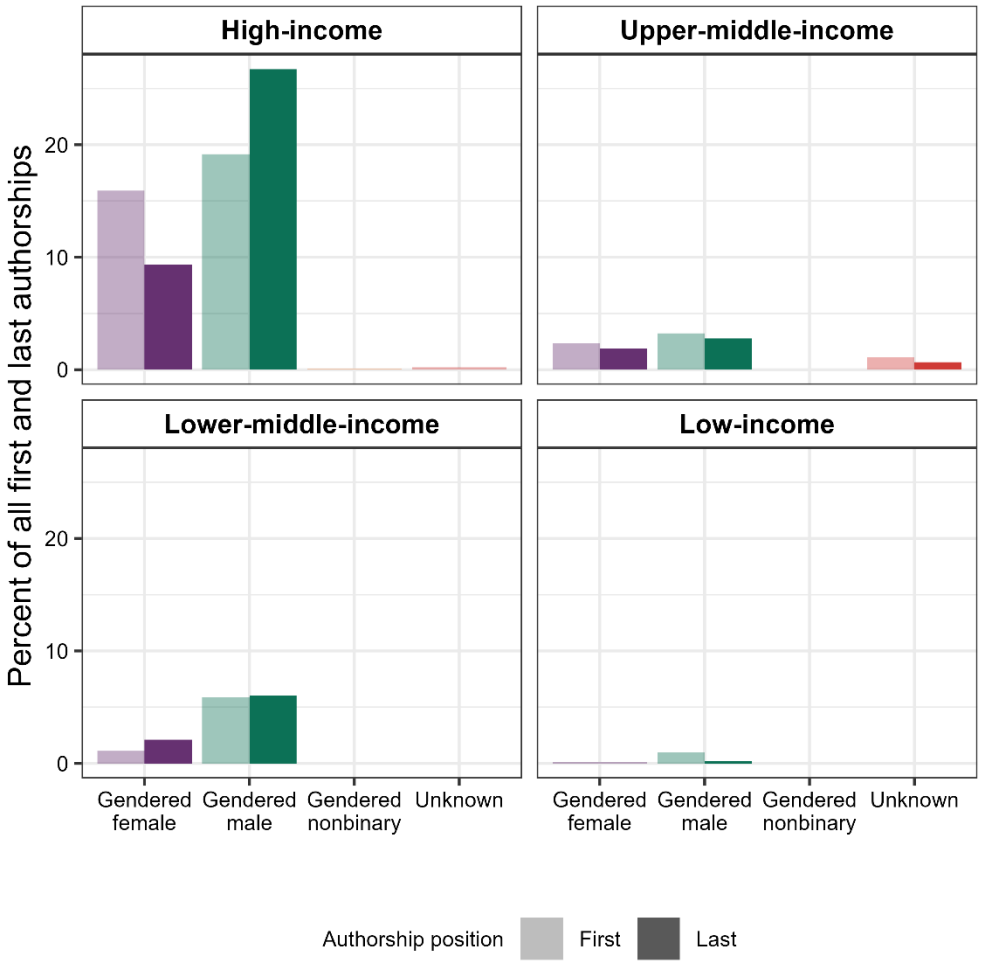


Fig 2. Percent of all first and last authorships (*n* = 898) separated by country income. Colors indicate author gender and shading indicates authorship position. A total of 43 countries were represented in first and last authorship affiliations (high-income: *n* = 19, upper-middle-income: *n* = 8, lower-middle-income: *n* = 13, low-income: *n* = 3).

Gender composition of authorships over time

The percent of first authorships by gendered female authors fluctuated from year to year, displaying no clear trend over time (Fig. 3A). The longest period of monotonic growth was from 2016 to 2019, when first authorships by gendered female authors increased from 26.7% to

48.9%; this was followed by a monotonic decline to 30.0% by 2022. Despite the fluctuations, the percent of first authorships by gendered female authors never dipped below 25% nor reached above 50% over the 12-year timespan of our dataset. In contrast, there appeared to be a positive trend over time for last authorships by gendered female authors (Fig. 3A). In 2011, there were zero last authorships by gendered female authors. This value jumped to 14.3% in 2012, remained fairly constant through 2015, and jumped again to 39.3% in 2016. There was a subsequent monotonic decline to 17.9% by 2020. The percent of last authorships jumped to a maximum of 43.9% in 2021, then decreased to 25.0% in 2022.

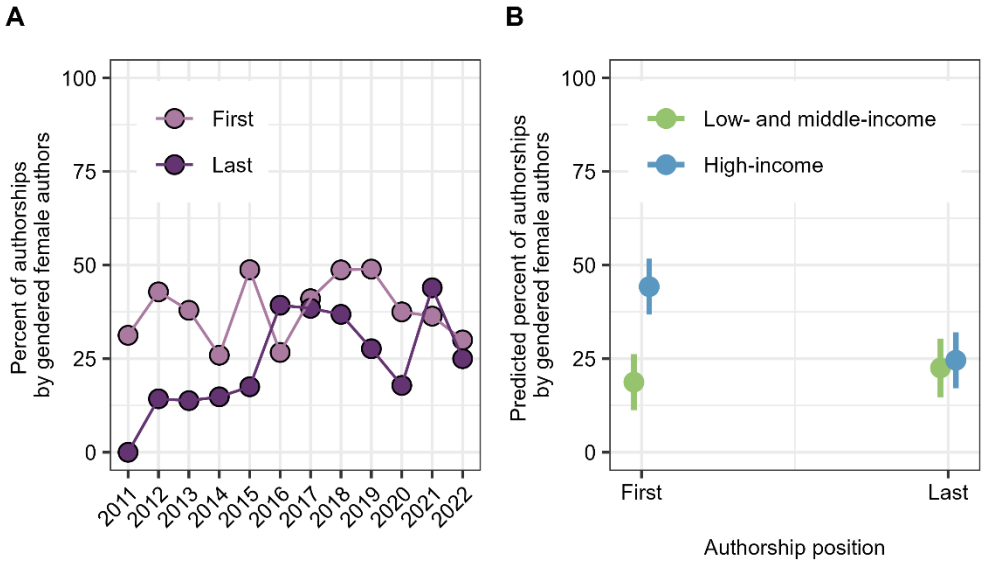


Fig 3. Observed and model-predicted gendered female authorship. (A) Observed percent of authorships by gendered female authors over time, separated by authorship position. (B) Model-predicted percent of authorships by gendered female authors, separated by authorship position and country income.

A linear model explaining the percent of authorships by gendered female authors found that the interaction between year and authorship position was not statistically significant (β = 1.45, *SE* = 1.11, *P* = 0.20; Table S3). The main effect of year was positive, indicating an overall increase in authorships by gendered female authors, though this was also not statistically significant (β = 0.85, *SE* = 0.76, *P* = 0.27). There was a significant interaction between authorship position and country income (β = -23.45, *SE* = 7.49, *P* = 0.0032), meaning that the relationship between authorship position and the percent of authorships by gendered female authors depended on country income. For high-income countries, predicted authorships by gendered female authors were greater for first authorships (44.3%, 95% confidence interval: 36.8-51.7) compared to last authorships (24.6%, 95% CI: 17.1-32.1; Fig. 3B). However, for low- and middle-income countries, the predicted first (18.74%, 95% CI: 11.3-26.2) and last (22.5%, 95% CI: 14.7-30.4) authorships by gendered female authors were similar. Accounting for the number of variables, the model explained 41% of the variance in the outcome variable.

Alignment of authorship geography and article geography

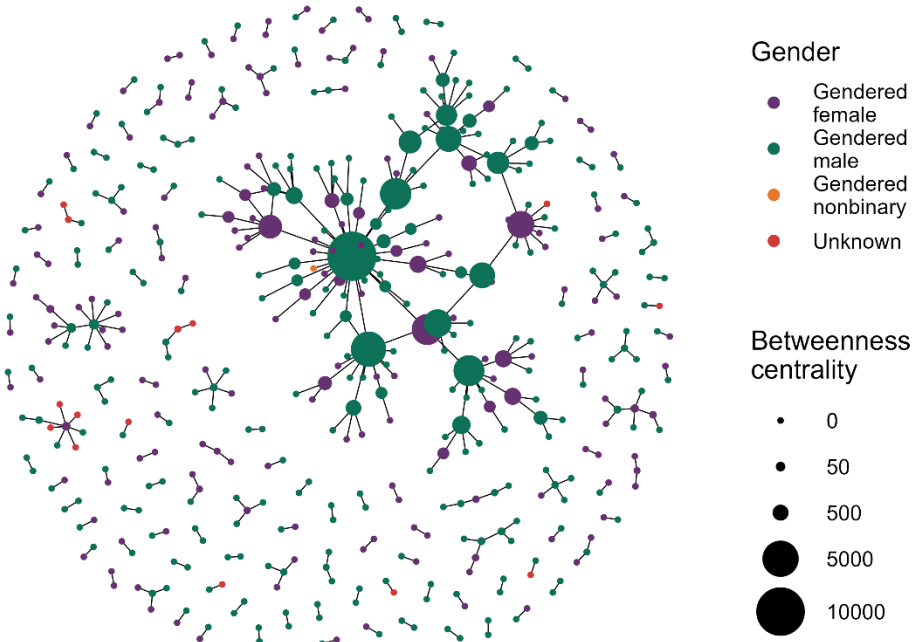
We found that 280 articles (62.1% of all articles in the corpus) were geographically focused on one or more non-US countries (e.g. where field or laboratory work was performed or specialized knowledge of a country was required). About two-thirds of the time (63.9%, 179/280), either the first or last authorship geography matched the article geography (Table S4). First authorship geography matched article geography 53.6% of the time, whereas last authorship geography matched article geography 47.1% of the time. Both first and last authorships matched the article geography 36.7% of the time. Though there were more articles published in the second half of the dataset (2011-2016: 95 articles; 2017-2022: 185 articles), the rates of

386 authorship-geography matches stayed nearly constant over time (Table S4). There were 142
387 articles (31.5%) with a non-specific article geography and 29 articles (6.4%) with an article
388 geography that included the United States.
389

390 ## Characteristics of the author network

391 We created two author network depictions to show relationships between author gender
392 and our two centrality measures of interest (Fig. 4). The mean betweenness centrality of the
393 gendered male author group was almost double that of the gendered female author group (129.21
394 versus 64.84; Table S5), indicating that on average, gendered male authors were more
395 structurally important to the network. There were nearly twice as many gendered male authors in
396 positions of "power" in the network (betweenness centrality ≥ 317, the 95th percentile)
397 compared to gendered female authors (17 versus 9). In contrast, the two author groups had
398 similar mean harmonic centrality (gendered male: 11.51, gendered female: 11.25; Table S5),
399 indicating similar levels of connection or collaboration. There were 12 gendered female authors,
400 15 gendered male authors, and one nonbinary author who were highly collaborative (harmonic
401 centrality ≥ 43.1, the 95th percentile).
402

A


Gender
Gendered female
Gendered male
Gendered nonbinary
Unknown
Betweenness centrality
0
50
500
5000
10000

B

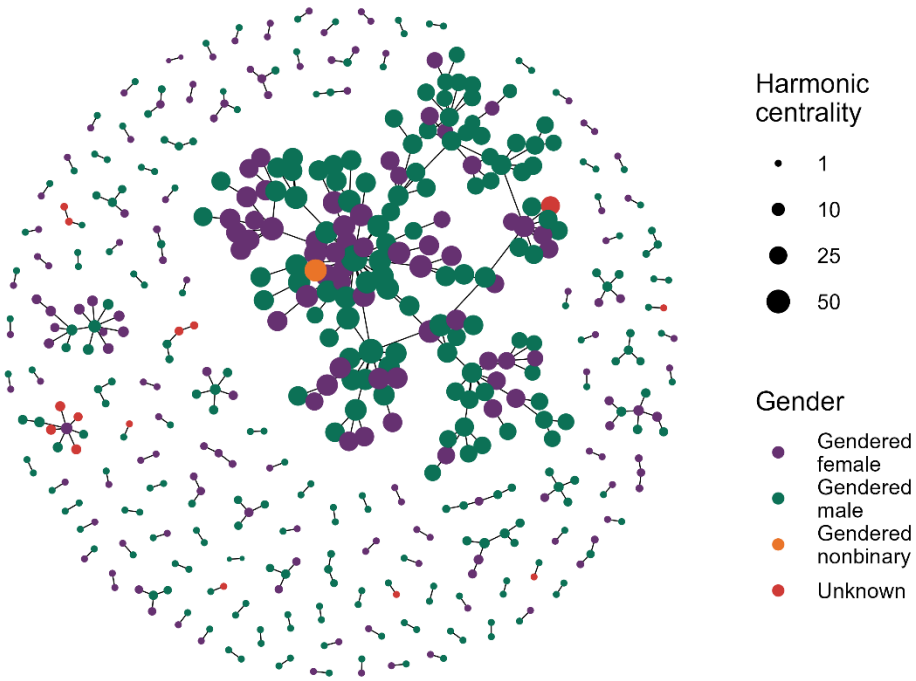

Harmonic centrality
1
10
25
50
Gender
Gendered female
Gendered male
Gendered nonbinary
Unknown

Fig 4. Depictions of the author network by two measures of centrality. In both panels, node (author) color indicates gender. (A) Network with node size representing betweenness centrality, which conveys information about the structural importance of an individual to the overall network. (B) Network with node size representing harmonic centrality, which conveys information about how connected an individual is to the rest of the graph.

Discussion

We analyzed authorship patterns at EcoHealth Alliance, a United States-based organization that conducts One Health and conservation research. Within our corpus of 451 peer-reviewed journal articles, we found that gendered male authors were dominant in multiple aspects: they represented ~60% of all authors, 65% of all FLAs, and 91% of highly productive authors. The last authorship position was particularly male-dominated, with 2.66 times as many last authorships by gendered male authors as by gendered female authors. Gendered male authors were more structurally important to the author network on average and comprised 65% of highly "powerful" authors in the network. We also found overrepresentation of authorships affiliated with high-income countries, with ~72% of FLAs listing a high-income country affiliation. The over-representation of gendered male and high-income country authors in prestigious authorship positions shows publication practices at EHA could be changed to improve equity.

We focused our analyses on first and last authorships because these are both typically perceived as prestigious in STEM fields, though for different reasons. In the ecological and environmental sciences, the first author is commonly viewed as taking the lead role in study conceptualization, data collection and analysis, and manuscript writing, whereas the last author is often viewed as the "senior" author whose work or role made the study possible (40-43). Our

finding of more first than last authorships for gendered female authors, but the reverse pattern for gendered male authors, suggests that gendered female authors were more likely to lead papers whereas gendered male authors were more likely to lead research groups. This is supported by our finding that the most productive authors in our dataset (i.e. those with ≥ 10 total FLA) were nearly all gendered male (10/11), with last authorships comprising a large percentage of their FLA. Leading a publication as a first author is a time-consuming endeavor that requires intensive analysis and writing; therefore, the overrepresentation of gendered female authors in this role may result in decreased overall productivity. In contrast, research supervisors or principal investigators can accumulate last authorships on their team members' publications for comparatively less effort per publication (e.g. general oversight and manuscript editing). It is important to note that the overall authorship pattern we observed (more first than last authorships for gendered female authors, but more last than first authorships for gendered male authors) was driven primarily by authorships with high-income country affiliations. Further work focusing on the interplay between authorship position and gender for researchers in low- and middle-income countries would be especially valuable.

Our analysis of the authorship network shows that gendered male authors were disproportionately represented in structurally important positions based on betweenness centrality scores. However, gendered male authors had similar collaborativeness as gendered female authors: there were similar numbers of gendered male and gendered female authors in the top 5th percentile of harmonic centrality scores, and average harmonic centrality scores for the two gender groups were nearly identical. Together, these results show that gendered female authors were just as collaborative as their gendered male peers, but it was gendered male authors

who were more likely to be in positions of power (i.e. those who control or distribute resources like funding).

Though EcoHealth Alliance primarily conducts research in low- and middle-income countries and aims to engage local partners, we found that credit in the form of prestigious first and last authorships went to authors affiliated with the United States or another high-income country more than half of the time. When we examined articles that were geographically focused outside the United States—representing an opportunity for local leadership—first and last authorship geography each matched the article geography about half the time. In ~36% of articles, neither the first nor the last authorship geography matched the article geography. These findings echo previous studies showing that when research takes place in or involves participants from a low- or middle-income country, researchers affiliated with that country are only included as first or last authors about half the time (44-46). One reason for this disparity could be the devaluation of certain steps in the scientific process (e.g. project implementation, data collection) that are usually conducted by low-and middle income country researchers in comparison to others typically conducted by high-income country researchers (e.g. drafting a manuscript, acquiring research funding) (47). Concerns about the potential for editorial bias, where journals may favor well-known authors or those from English-speaking countries, might also lead low- and middle income country researchers to cede first or last authorship to high-income country collaborators to increase the chance that an article will be accepted for publication (47). Authors from high-income countries may prioritize first or last authorships if they believe this will improve their chances of securing future research funding (6).

Contrary to our expectations, neither the interaction between authorship position and year, nor the main effect of year, were statistically significant variables in our linear model

explaining the percent of authorships by gendered female authors. However, the interaction between country income and authorship position was an important explanatory variable. For high-income countries, there was a striking disconnect in the model-predicted percent of first authorships (~44%) versus last authorships (~25%) by gendered female authors (Fig. 3B). This could represent a lack of career advancement for gendered female authors, who may publish primarily first authorships as early or mid-career researchers but rarely become senior researchers with a shift to publishing primarily last authorships.

Conclusion

Our results suggest that publication practices at EcoHealth Alliance favored gendered male researchers and researchers from high-income countries in prestigious authorship positions. Though our analysis was based on a corpus of articles with at least one EHA-affiliated author, publication practices at EHA are likely not unique among One Health and conservation research organizations based in high-income countries (15). To begin to address this disparity, scientists affiliated with organizations in high-income countries should not impose their own authorship norms when collaborating with peers from low- or middle-income countries, and should make an effort to familiarize themselves with the norms of their collaborators. We recommend having early (and ongoing) conversations about authorship practices and expectations with the entire research team. Adopting a consensus-based decision making process and considering the different types of labor involved in a project, who performed that labor, and the "social location" of each author may be useful strategies for determining authorship status and order (48).

We recommend that research organizations in high-income countries critically evaluate how organizational structure (including who is allowed or encouraged to be a principal

investigator) can impact scientists' publication records. To avoid individuals in positions of power imposing their own authorship norms, organizations should develop guidelines around authorship order that recognize the different types of contributions that scientists make (49). When organizations in high-income countries conduct research that requires collaboration with individuals in low- or middle-income countries, these authorship guidelines should explicitly require inclusion of authors from those partnering countries in first and last authorship positions (50). To incentivize more equitable authorship allocations, organizations could consider publications that list collaborators from low- or middle-income countries (especially those who are gendered female) in first and last authorship positions as a strength in performance evaluations and decisions around promotion.

Finally, as gatekeepers at different points in the project life cycle, funders and publishers ultimately control what research is conducted, how it is conducted, and how findings are distributed. As such, they should take a more active role in requiring researchers to reflect on and acknowledge contributions from all collaborators (51). For example, the journal *BMJ Global Health* requires authors to provide a structured reflexivity statement (50) when submitting manuscripts involving collaboration between researchers from high-income and low- or middle-income countries. When evaluating a project proposal, funders should require equitable allocation of intellectual property and scholarly recognition between researchers of high and low- or middle-income countries. Funders should also account for structural biases that may have shaped applicants' publication records when evaluating researchers' capacity to conduct a project. Taken together, these actions by funders and publishers would create top-down pressure on researchers and organizations to improve equity in how research is conducted and published.

We hope our results will inspire changes in international research collaborations and authorship practices at EcoHealth Alliance and similar organizations. Individual researchers, research organizations, funders, and publishers all play an important role in ensuring that acknowledgement of scholarly contributions is more equitable, and that marginalized scientists receive the recognition that they deserve.

Acknowledgements

We thank EcoHealth Alliance for allowing us to study publication practices at the organization. We are grateful to the WiSE group at EHA for their general support and for stimulating discussions around gender and science. Thank you to members of the Carlson Lab at Yale University for providing constructive feedback on a previous version of this manuscript.

References

1. Bennett DM, Taylor DM. Unethical practices in authorship of scientific papers. Emergency Medicine. 2003;15(3):263-70.

2. Moher D, Naudet F, Cristea IA, Miedema F, Ioannidis JPA, Goodman SN. Assessing scientists for hiring, promotion, and tenure. PLoS Biol. 2018;16(3):e2004089.

3. Niles MT, Schimanski LA, McKiernan EC, Alperin JP. Why we publish where we do: Faculty publishing values and their relationship to review, promotion and tenure expectations. PLOS ONE. 2020;15(3):e0228914.

4. Rice DB, Raffoul H, Ioannidis JPA, Moher D. Academic criteria for promotion and tenure in biomedical sciences faculties: cross sectional analysis of international sample of universities. BMJ (Clinical research ed). 2020;369:m2081.

5. Ebadi A, Schiffauerova A. How to Receive More Funding for Your Research? Get Connected to the Right People! PLOS ONE. 2015;10(7):e0133061.

6. Michelson KA. Association of publication record and independent NIH funding. PLOS ONE. 2022;17(6):e0269283.

7. Samaniego C, Lindner P, Kazmi MA, Dirr BA, Kong DT, Jeff-Eke E, et al. Higher research productivity = more pay? Gender pay-for-productivity inequity across disciplines. Scientometrics. 2023;128(2):1395-407.

8. Wren JD, Kozak KZ, Johnson KR, Deakyne SJ, Schilling LM, Dellavalle RP. The write position. EMBO reports. 2007;8(11):988-91-91.

9. Larivière V, Desrochers N, Macaluso B, Mongeon P, Paul-Hus A, Sugimoto CR. Contributorship and division of labor in knowledge production. Social Studies of Science. 2016;46(3):417-35.

10. von Bartheld CS, Houmanfar R, Candido A. Prediction of junior faculty success in biomedical research: comparison of metrics and effects of mentoring programs. PeerJ. 2015;3:e1262.

11. Wapman KH, Zhang S, Clauset A, Larremore DB. Quantifying hierarchy and dynamics in US faculty hiring and retention. Nature. 2022;610(7930):120-7.

12. National Center for Science and Engineering Statistics (NCSES). Diversity and STEM: Women, Minorities, and Persons with Disabilities 2023. Alexandria, VA: National Science Foundation; 2023.

13. Larivière V, Ni C, Gingras Y, Cronin B, Sugimoto CR. Bibliometrics: Global gender disparities in science. Nature. 2013;504(7479):211-3.

14. West JD, Jacquet J, King MM, Correll SJ, Bergstrom CT. The Role of Gender in Scholarly Authorship. PLOS ONE. 2013;8(7):e66212.

15. James R, Ariunbaatar J, Bresnahan M, Carlos-Grotjahn C, Fisher JRB, Gibbs B, et al. Gender and conservation science: Men continue to out-publish women at the world's largest environmental conservation non-profit organization. Conservation Science and Practice. 2022;4(8):e12748.

16. Holman L, Stuart-Fox D, Hauser CE. The gender gap in science: How long until women are equally represented? PLoS Biol. 2018;16(4):e2004956.

17. King MM, Frederickson ME. The Pandemic Penalty: The Gendered Effects of COVID-19 on Scientific Productivity. Socius. 2021;7:23780231211006977.

18. Ni C, Smith E, Yuan H, Larivière V, Sugimoto CR. The gendered nature of authorship. Science Advances. 2021;7(36):eabe4639.

19. Ross MB, Glennon BM, Murciano-Goroff R, Berkes EG, Weinberg BA, Lane JI. Women are credited less in science than men. Nature. 2022;608(7921):135-45.

20. Hedt-Gauthier BL, Jeufack HM, Neufeld NH, Alem A, Sauer S, Odhiambo J, et al. Stuck in the middle: a systematic review of authorship in collaborative health research in Africa, 2014-2016. BMJ Global Health. 2019;4(5):e001853.

21. Dimitris MC, Gittings M, King NB. How global is global health research? A large-scale analysis of trends in authorship. BMJ Global Health. 2021;6(1):e003758.

22. Shambe I, Thomas K, Bradley J, Marchant T, Weiss HA, Webb EL. Bibliometric analysis of authorship patterns in publications from a research group at the London School of Hygiene & Tropical Medicine, 2016-2020. BMJ Global Health. 2023;8(2):e011053.

23. Priem J, Piwowar H, Orr R. OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. arXiv. 2022:2205.01833.

24. Aria M, Le T, Cuccurullo C, Belfiore A, Choe J. openalexR: An R-Tool for Collecting Bibliometric Data from OpenAlex. The R Journal. 2024;15(4):167-80.

25. Sánchez C, Schwantes C, Ball S, Kim S, Munro S, Leaman B, et al. Database of author demographics of scientific publications at EcoHealth Alliance. 0.0.1 ed: Zenodo; 2024.

26. Conlin J. The Freedom to Choose Your Pronoun. The New York Times. 2011 October 2.

27. American Dialect Society. 2019 Word of the Year is "(My) Pronouns," Word of the Decade is Singular "They": Duke University Press; 2020 [Available from: https://americandialect.org/2019-word-of-the-year-is-my-pronouns-word-of-the-decade-is-singular-they/].

28. Zettlemoyer MA, Cortijo-Robles KM, Srodes N, Johnson SE. What are we reading? Hot Topics and Authorship in Ecology Literature Across Decades. The Bulletin of the Ecological Society of America. 2023;104(1):e02025.

29. Jiang J, Chen E, Luceri L, Muric G, Pierri F, Chang H-CH, et al. What are Your Pronouns? Examining Gender Pronoun Usage on Twitter. arXiv. 2023.

30. Stonewall UK. List of LGBTQ+ terms 2024 [Available from: https://www.stonewall.org.uk/resources/list-lgbtq-terms].

31. Van Buskirk I, Clauset A, Larremore DB. An Open-Source Cultural Consensus Approach to Name-Based Gender Classification. Proceedings of the International AAAI Conference on Web and Social Media. 2023;17(1):866-77.

32. Frances DN, Fitzpatrick CR, Koprivnikar J, McCauley SJ. Effects of Inferred Gender on Patterns of Co-Authorship in Ecology and Evolutionary Biology Publications. The Bulletin of the Ecological Society of America. 2020;101(3):e01705.

33. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2024.

34. Newman M. Networks: Oxford University Press; 2018. [Available from: https://doi.org/10.1093/oso/9780198805090.001.0001].

35. Yan E, Ding Y. Applying centrality measures to impact analysis: A coauthorship network analysis. Journal of the American Society for Information Science and Technology. 2009;60(10):2107-18.

36. Freeman LC. A Set of Measures of Centrality Based on Betweenness. Sociometry. 1977;40(1):35-41.

37. Marchiori M, Latora V. Harmony in the small-world. Physica A: Statistical Mechanics and its Applications. 2000;285(3):539-46.

38. Csardi G, Nepusz T. The igraph software package for complex network research. InterJournal. 2006;Complex Systems:1695.

39. Csárdi G, Nepusz T, Müller K, Horvát S, Traag V, Zanini F, et al. igraph for R: R interface of the igraph library for graph theory and network analysis (v2.1.4). Zenodo; 2025.

40. Weltzin JF, Belote RT, Williams LT, Keller JK, Engel EC. Authorship in ecology: attribution, accountability, and responsibility. Front Ecol Environ. 2006;4(8):435-41.

41. Tscharntke T, Hochberg ME, Rand TA, Resh VH, Krauss J. Author Sequence and Credit for Contributions in Multiauthored Publications. PLoS Biol. 2007;5(1):e18.
42. Duffy MA. Last and corresponding authorship practices in ecology. Ecology and Evolution. 2017;7(21):8876-87.
43. Logan JM, Bean SB, Myers AE. Author contributions to ecological publications: What does it mean to be an author in modern ecological research? PLOS ONE. 2017;12(6):e0179956.
44. Kelaher M, Ng L, Knight K, Rahadi A. Equity in global health research in the new millennium: trends in first-authorship for randomized controlled trials among low- and middle-income country researchers 1990-2013. International Journal of Epidemiology. 2016;45(6):2174-83.
45. Mbaye R, Gebeyehu R, Hossmann S, Mbarga N, Bih-Neh E, Eteki L, et al. Who is telling the story? A systematic review of authorship for infectious disease research conducted in Africa, 1980â€ " 2016. BMJ Global Health. 2019;4(5):e001855.
46. Modlin CE, Deng Q, Benkeser D, Mulate YW, Aseffa A, Waller L, et al. Authorship trends in infectious diseases society of America affiliated journal articles conducted in low-income countries, 1998–2018. PLOS Global Public Health. 2022;2(6):e0000275.
47. Smith E, Hunt M, Master Z. Authorship ethics in global health research partnerships between researchers from low or middle income countries and high income countries. BMC Medical Ethics. 2014;15(1):42.
48. Liboiron M, Ammendolia J, Winsor K, Zahara A, Bradshaw H, Melvin J, et al. Equity in Author Order: A Feminist Laboratory's Approach. Catalyst: Feminism, Theory, Technoscience. 2017;3(2):1-17.

49. Miles S, Renedo A, Marston C. Reimagining authorship guidelines to promote equity in co-produced academic collaborations. Global Public Health. 2022;17(10):2547-59.

50. Morton B, Vercueil A, Masekela R, Heinz E, Reimer L, Saleh S, et al. Consensus statement on measures to promote equitable authorship in the publication of research from international partnerships. Anaesthesia. 2022;77(3):264-76.

51. Kiermer V. Authorship practices must evolve to support collaboration and open science. PLoS Biol. 2023;21(10):e3002364.

Supporting information

Fig S1. Percent of all first and last authorships (n = 898) separated by author gender and authorship position.

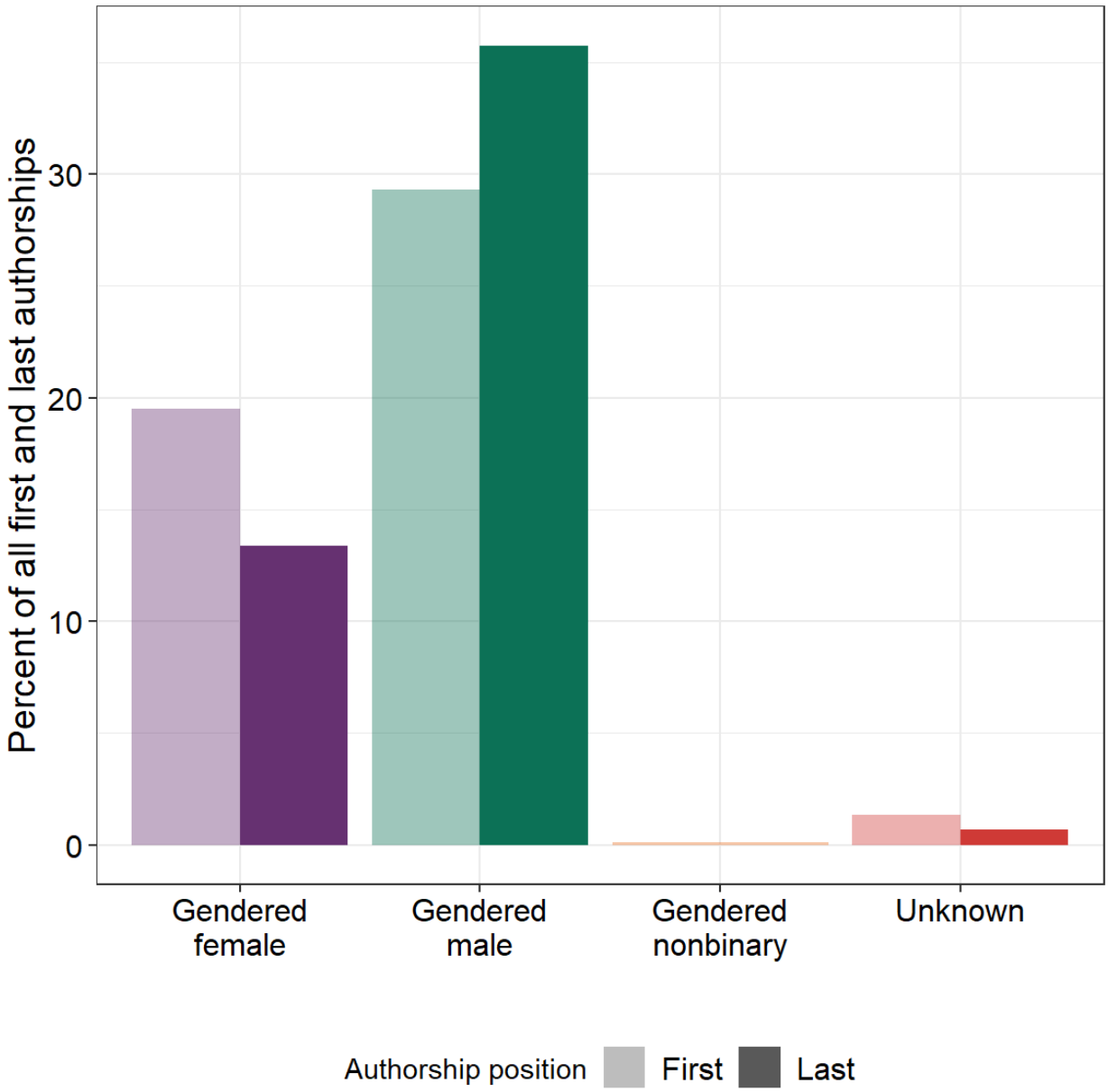


Fig S2. Number of first and last authorships (n = 898) separated by country affiliation.

Points are colored according to country income. Note that the x-axis is on a log-10 scale.

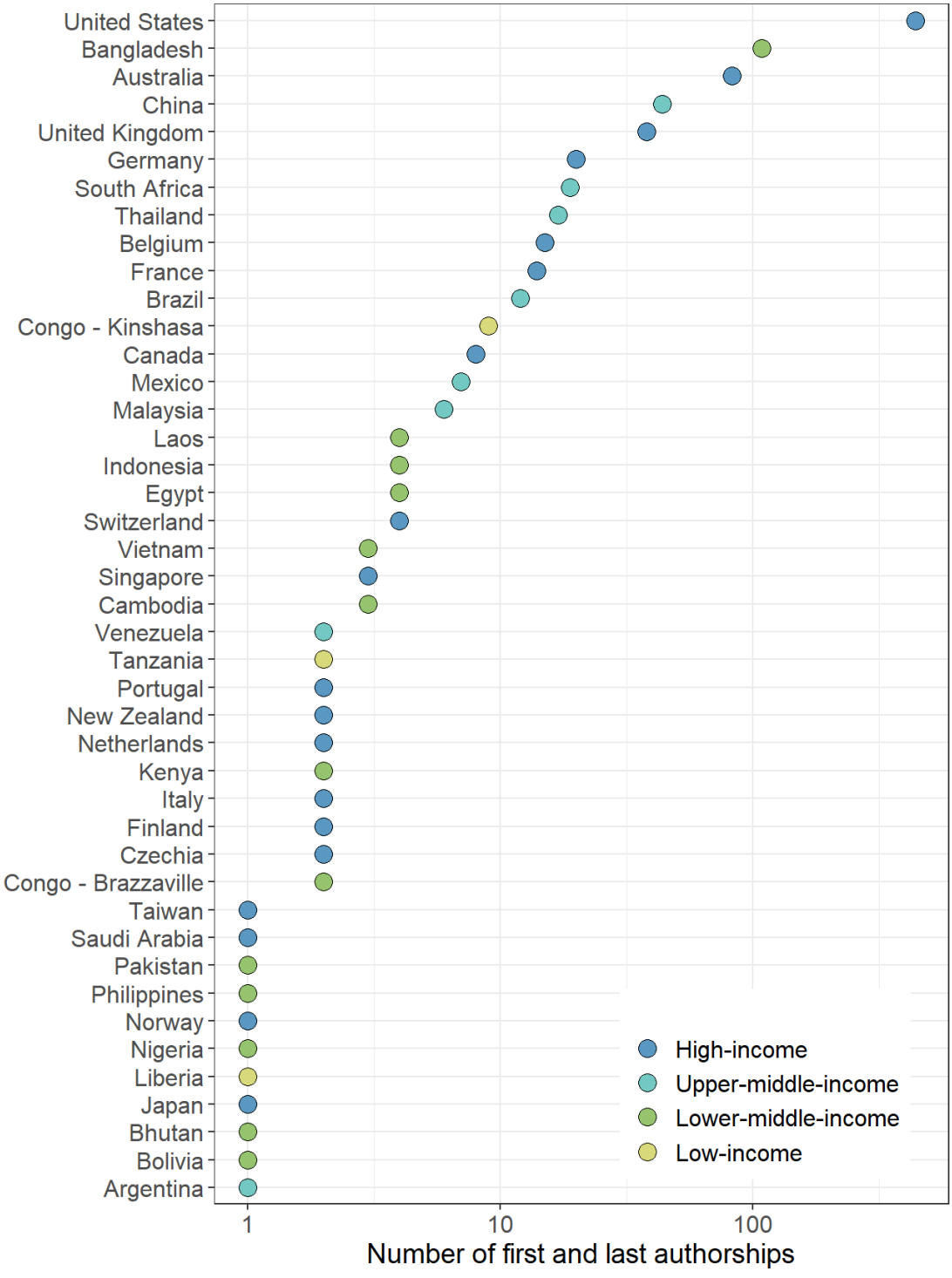


Fig S3. Number of first and last authorships (*n* = 898) by country affiliation and gender. Data for the five countries with the most first and last authorships are displayed individually, while data for the remaining countries (*n* = 37) are grouped into "Other".

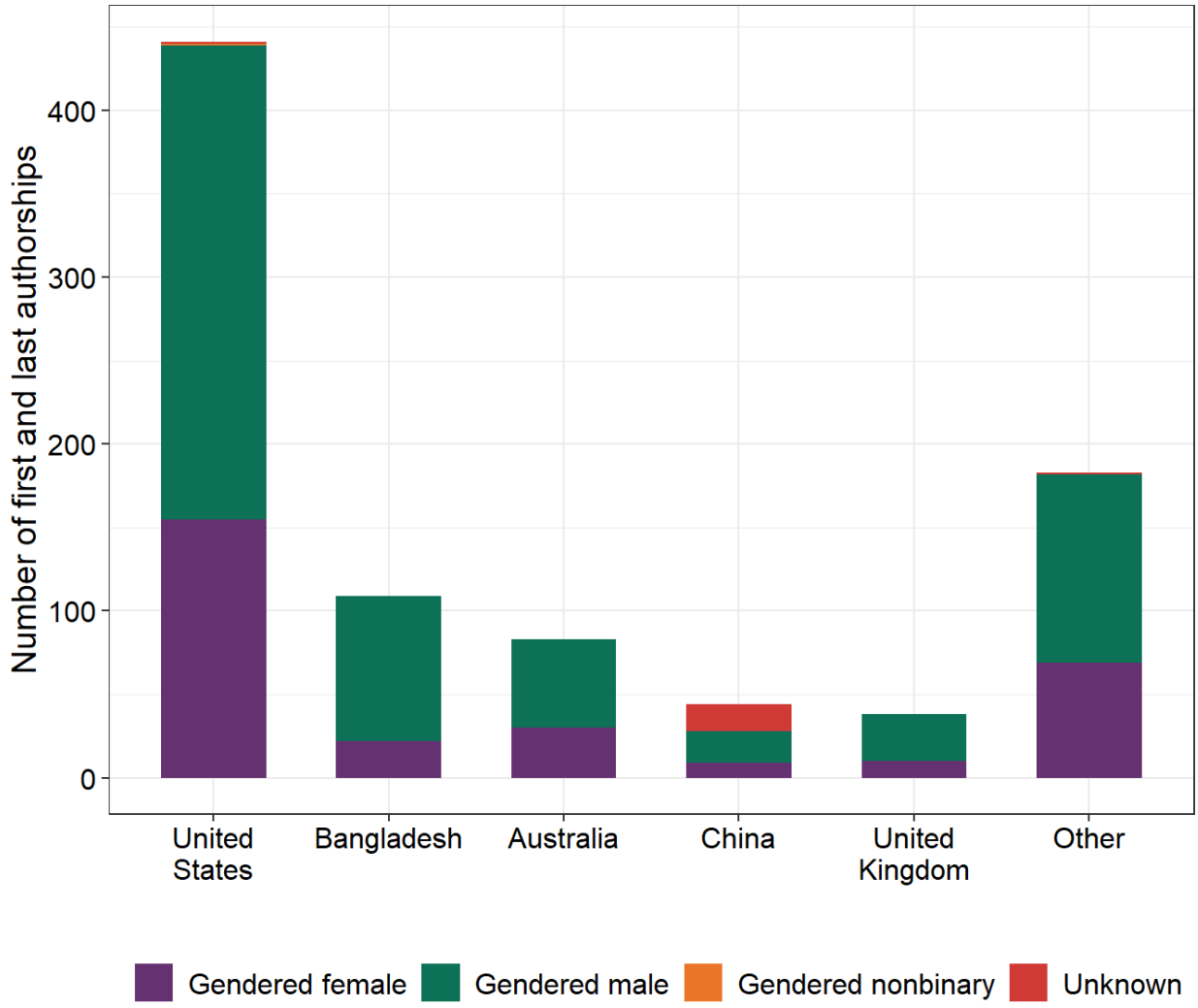


Fig S4. Percent of all first and last authorships (n = 898) separated by country income and authorship position.

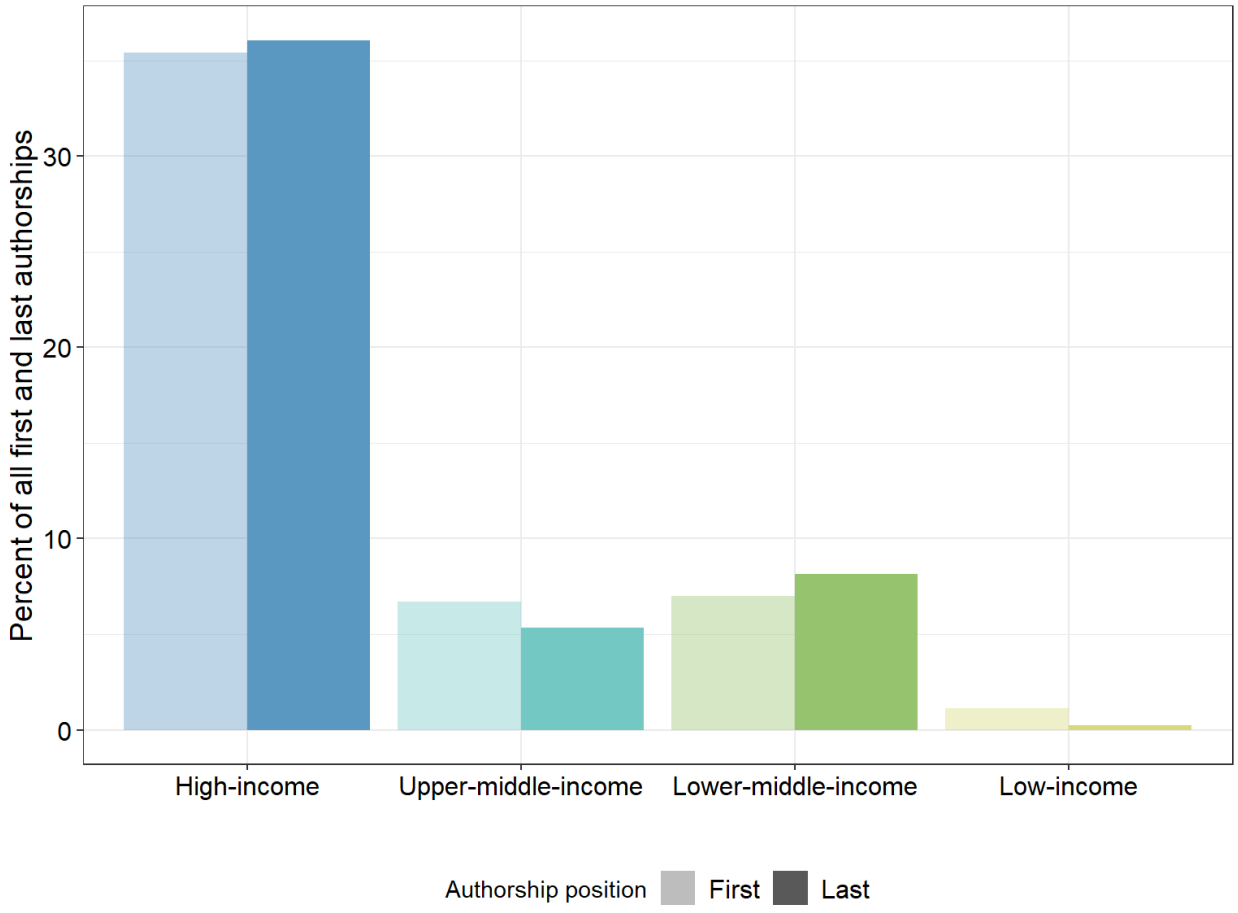


Table S1. Comparison of a pronouns-based approach and a name-based approach to classify author genders. See *Gender classification of authors* in the main text for details of how authors were classified using each approach.

		Gender classification based on name			
		gendered female	*gendered male*	*uncertain*	*undetermined*
Gender classification based on pronouns	*gendered female*	151	3	20	7
	gendered male	3	244	24	9
	gendered nonbinary	0	1	0	0
	undetermined	5	17	12	2

Table S2. Last authorships as a percentage of all first and last authorships (FLAs) for the most productive authors in the dataset (i.e. ≥ 10 FLAs).

Total FLAs	Author gender	Last authorships/total FLAs (%)
44	Gendered male	88.6
30	Gendered male	43.3
22	Gendered male	68.2
18	Gendered male	77.8
15	Gendered male	53.3
14	Gendered male	92.9
12	Gendered female	100
12	Gendered male	100
11	Gendered male	18.2
11	Gendered male	72.7
10	Gendered male	70

Table S3. Model coefficients for a linear model to examine effects of authorship position, country income, and year on the percent of authorships by gendered female authors. The "Year" variable was centered around 2011 to improve coefficient interpretability. *P* values < 0.05 are bolded.

Variable	**Estimate**	**SE**	**t**	***P***
Intercept	13.97	5.57	2.51	**0.016**
Year	0.85	0.76	1.12	0.27
Authorship position(Last)	-4.35	8.33	-0.52	0.60
Country income(High)	25.51	5.23	4.87	**1.7e-5**
Year : Authorship position(Last)	1.45	1.11	1.31	0.20
Authorship position(Last) : Country income(High)	-23.45	7.49	-3.13	**0.0032**

Table S4. Matches between authorship geography (i.e. country affiliation) and article geography (i.e. the geographic focus of a study, excluding the United States). A "match" occurred when the authorship geography was the same as any of the countries contained in the article geography. Values are provided for the whole timespan of the data (2011-2022) as well as broken down into two time periods (2011-2016 and 2017-2022) to explore potential changes in authorship over time. Denominator sizes are sometimes different because not all articles had a last authorship (i.e. sole-authored articles, which were counted as first authorships).

Time period	First authorship match	Last authorship match	Either match	Both match
2011-2022	150/280 (53.6%)	131/278 (47.1%)	179/280 (63.9%)	102/278 (36.7%)
2011-2016	51/95 (53.7%)	43/94 (45.7%)	61/95 (64.2%)	33/94 (35.1%)
2017-2022	99/185 (53.5%)	88/184 (47.8%)	118/185 (63.8%)	69/184 (37.5%)

Table S5. A summary of two measures of network centrality (betweenness centrality and harmonic centrality) calculated for authors separated by gender. HDCI = high density confidence interval.

Gender	Group size	Mean betweenness centrality	Betweenness centrality 95% HDCI	Mean harmonic centrality	Harmonic centrality 95% HDCI
Gendered female	186	64.84	0-190.6	11.25	0.65-43.12
Gendered male	297	129.21	0-317.0	11.51	0.83-43.12
Gendered nonbinary	1	0	–	43.12	--
Unknown	14	0.21	0-2.0	3.95	1.00-28.51