# Persistent Hierarchy in Contemporary International Collaboration

**Abstract**

Science has become increasingly globalized, with international collaboration playing a crucial role in advancing scientific development and facilitating knowledge exchange across borders. However, contemporary international collaboration is still influenced by power asymmetries rooted in colonial legacies, resulting in researchers from non-Western countries often marginalized. This study investigates this power structure in international collaboration. We examine the underlying factors shaping the power structures in these collaborations and the consequences of such asymmetries. Our findings reveal that international collaboration is embedded in a hierarchical structure, where researchers from scientifically advanced countries are overrepresented in dominant roles while researchers from lagging countries are relegated to supportive roles. This power asymmetry also affects the research direction of international collaborative publications. As the scientific capacity of a country decreases, the content similarity between internationally coauthored publications and those without international collaboration also decreases. By analyzing the power dynamics within international collaborations, we reveal that researchers from less-developed countries are disadvantaged in the collaborative process which adversely impact both the global scientific community and national scientific development.

## Introduction

International collaboration is increasingly become the cornerstone of scientific progress, especially in addressing complex global challenges like vaccine development during pandemics and combating climate change—problems that demand the participation of researchers across the globe. However, while “collaboration” is often idealized as an equal partnership, the current

international collaborative framework is steeped in inequities and exploitation. Shaped largely by the ideologies and practices of Western countries, the system systematically supports and prioritizes the research agendas and careers of scientists from these nations[1–4]. In contrast, researchers from less developed countries are often marginalized, underrepresented in prominent authorship positions[5,6] or even excluded entirely from co-authorship[7,8]. Their research agendas are frequently sidelined[9], and the science conducted in their home countries is detached from local societal needs[10].

This asymmetry, in which international collaboration centers around the interests of Western countries while researchers from non-Western countries occupy subordinate roles, reflects what world-systems theory describes as the core-periphery structure[11,12]. This theory argues despite the increasing global participation in science, international collaborations remain largely asymmetrical, with a few dominant nations—such as the U.S. and the UK—reproducing their scientific hegemony by maintaining the subordinate status of peripheral countries[13]. The position of individual researchers in this core-periphery hierarchy is predetermined by factors such as nationality, scientific heritage, and access to research infrastructures[14]. The structure is deeply entrenched in modern scientific institutions, largely due to the historical legacy of colonialism, where Western science advanced through the exploitation of resources and human capital in colonized regions[15,16]. During the colonial era, scientists from developed countries often exploited local resources and conducted research with the assistance of indigenous communities, yet rarely gave proper recognition to their contributions[16]. This practice cemented the ongoing marginalization and exploitation of non-Western researchers in contemporary international collaborations. Today, many developing countries not only depend on partnerships with developed nations to advance their scientific research but also rely on the material benefits these

collaborations provide, such as upgraded facilities and modernized laboratory infrastructure[17]. This dependency further strengthens the power imbalance, causing international collaboration to resemble the asymmetrical donor-recipient relationship, where developed countries hold the dominant position.

In response to this persistent power imbalance, there have been increasing calls to "decolonize science", advocating for a critical reevaluation and, in some cases, the dismantling of some modern scientific practices[18–20]. This landscape is further complicated by the rising of emerging scientific powerhouses, such as China and South Korea, which have begun to challenge the traditional core-periphery structure, leading to a more multilayered global scientific hierarchy[13,21–23]. Given the rising demand for global cooperation, alongside enduring inequities in international collaborations and the evolving landscape of the global scientific ecosystem, it is essential to examine the dynamics of international collaboration.

To address this knowledge gap, we leverage publication records from 201 countries to investigate the power dynamics in international collaborations globally. Each internationally coauthored publication is considered a product of a collaborative process, with the power dynamics inferred from the authorship. Since the order of authors on a paper typically indicates the level of contribution[24] and serve as a key feature for academic evaluation and promotion[25], we analyze these dynamics by examining the relationship between the authors' nationalities and their positions in the authorship list. Given that the first and last authors contribute most significantly to research[24,26] and receive the highest credit in evaluation systems[25], we classify authorship positions into leading roles (first and last authors) and supportive roles (middle

authors). We then investigate whether authors from less-advanced countries are disproportionately relegated to supportive roles, controlling for other influential factors, and explore the potential consequences of such discrimination.

**Results**

As the global scientific development and its integration march on, the number of authors participating in international collaborations has been steadily increasing, with researchers from scientifically advanced countries still comprising the largest proportion (see Methods, Fig. 1a-b). Meanwhile, there has been a substantial increase in the involvement of researchers from scientifically proficient countries in international collaborations, accompanied by a corresponding rise in the share (see Fig.1a). However, while the number of researchers from scientifically developing and lagging countries participating in international collaborations has also grown, their proportion remains modest (see Fig.1a). This trend holds across various authorship roles, with authors from scientifically advanced countries dominating the largest share across all positions, researchers from scientifically proficient countries have notably risen in all authorship positions, particularly as first authors, alongside a more modest contribution from researchers locate at developing and lagging countries (see Fig. 1b).

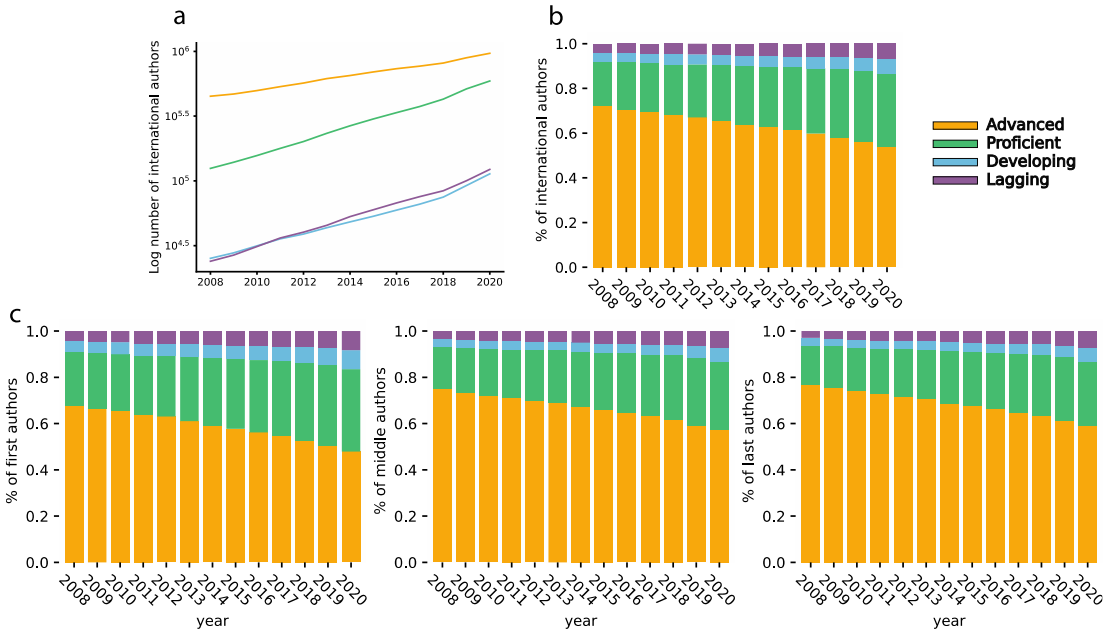



*Figure 1* ***Increasing proportion of authors from scientifically proficient countries in the international scientific workforce, especially as first authors.*** *(a) Number of authors participating in international collaborations within each scientific capacity group over years. (b) Proportion of international scientific workforce across different scientific capacity groups. (c) Proportion of authorships is occupied by authors from each scientific capacity group.*

Although authorship order typically reflects the importance and contribution of authors, the distribution of authorship positions is conditioned on the size of a country's scientific workforce. To assess whether researchers from less scientifically advanced countries are marginalized in key authorship positions, we normalize the raw occurrences in each authorship position of each country by the expected number (see Methods). The normalized results reveal a significant pattern: the authorship order in internationally coauthored publications follows a hierarchical structure. The role of the last author, typically associated with significant intellectual contribution and the greatest credit from the paper, is predominantly occupied by researchers from scientifically advanced countries. Notably, researchers from advanced countries are the only group overrepresented in last authorship, contrasting with researchers from non-advanced countries who are consistently underrepresented in this position (see Fig. 2a-b). Regarding the role of the first author, researchers from scientifically proficient countries exhibit a higher likelihood than expected of assuming this role, with a discernible upward trajectory wherein

researchers from scientifically developing countries are increasingly emerging as first authors in international collaborations. While there is a growing trend of researchers from scientifically proficient countries assuming the last authorship, they remain underrepresented in this role. Unfortunately, researchers from scientifically lagging countries are disproportionately underrepresented in both first and last authorship positions but overrepresented as middle authors compared to the expected value. Despite the majority of middle authors being situated in advanced countries, when accounting for the size of the scientific workforce, researchers from scientifically advanced countries exhibit a lower likelihood than expected of occupying the role of the middle author in international collaborations (see Fig. 2). However, given that the aggregated results might be influenced by publications authored by countries within the same scientific capacity group, we further conduct the same experiment excluding papers co-authored by countries from the same group. The results reveal consistent patterns, with a more pronounced divide between scientifically advanced and non-advanced countries (see Fig. S1). When researchers from scientifically advanced countries collaborate with those from different scientific capacity groups, they are significantly more likely to occupy the last author position (see Fig. S1).

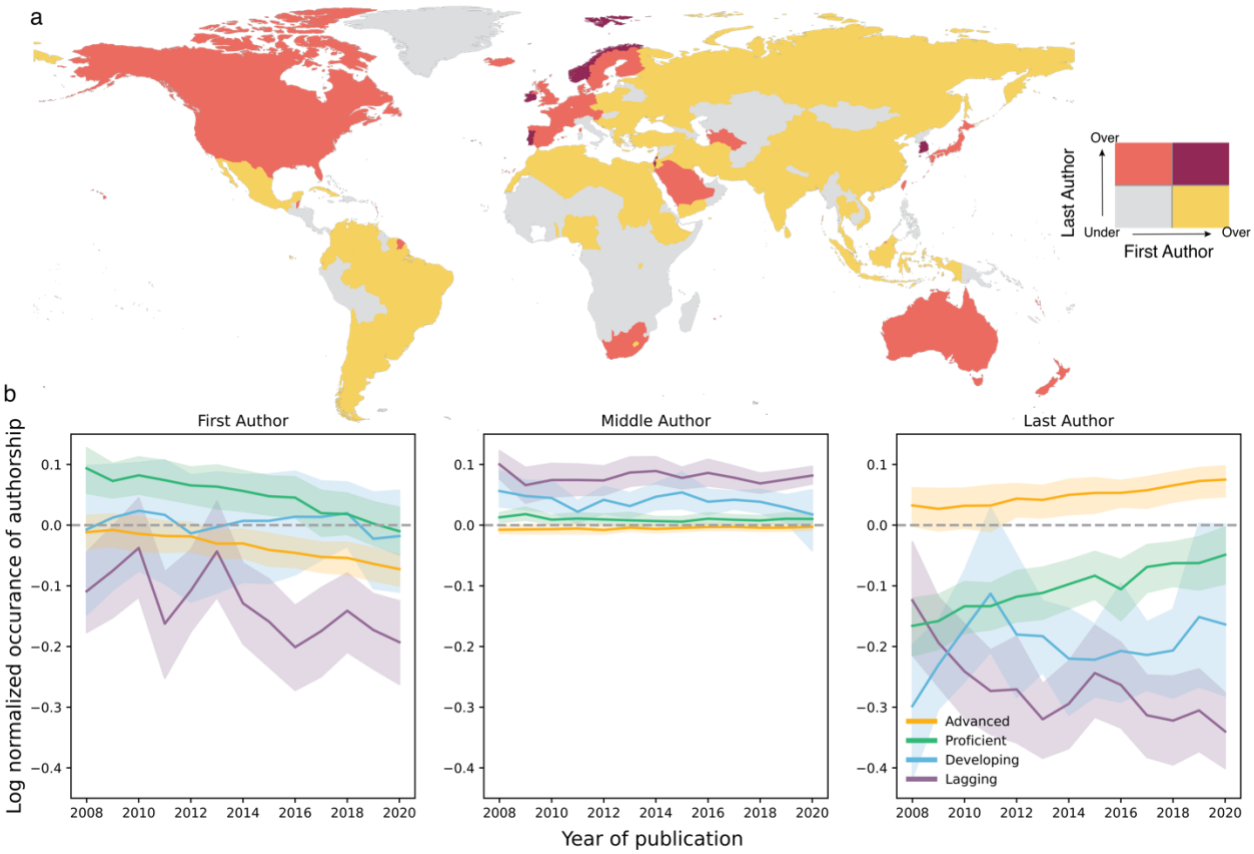
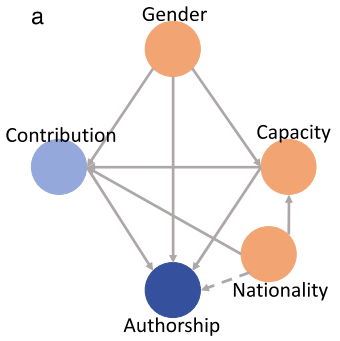



*Figure 2* ***The authorship order in internationally collaborated papers follows a hierarchcial structure.*** *(a) Countries are color-coded to indicate their representation in first and last authorship positions. Dark red represents that researchers from the focal country are overrepresented in both first and last authorship. Light red indicates researchers from the focal country are underrepresented in first authorship while overrepresented in last authorship. Yellow represents researchers from the country are overrepresented in first authorship but underrepresented in last authorship. The gray area represents researchers from the country are underrepresented in both first and last authorship. (b) Temporal trend. The occurrence of authorships by researchers from different countries is normalized against the expected number (see Methods). The Y-axis shows the logarithm of the normalized value. A positive value indicates that researchers from that country appear more frequently than expected, while a negative value indicates they appear less frequently than expected. The shaded area indicates the confidence interval, which is derived from aggregating countries in the same scientific group.*

While the normalization results reveal that researchers' role and importance in international collaboration is associated with their affiliation countries, the process of determining authorship order is simultaneously influenced by various factors, including the contribution of authors[24,26], or the author's gender[27] (see Data and Methods, see Fig. 3a). To identify the impact of nationality on authorship while controlling for the influence of other relevant factors, we apply a paper-level fixed-effects regression model (see Data and Methods). The regression results confirm that researchers from non-advanced countries tend to be systematically underrepresented

in the role of the last author; instead, they tend to assume the role of the first author and middle author (see Fig.3b). Specifically, women are more likely to serve as the first author than men, while less likely to serve as the last author. Compared to researchers from scientifically advanced countries, those from non-advanced countries are more likely to be the first author, instead of the last author, with researchers from scientifically proficient countries exhibiting the highest effect size. Meanwhile, researchers from non-advanced countries are also more likely to play the role of middle author, with the researchers from scientifically lagging countries have the highest effect size. While funding has a positive impact on increasing the likelihood of assuming leading roles as the first and last author, funded researchers from lagging countries still have a lower chance of serving as the last author compared to unfunded researchers from scientifically advanced countries (see Fig. 3c).

a




b

| | First Author | | | Middle Author | | | Last Author | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (1) | (2) | (3) | (1) | (2) | (3) |
| Male | -0.268*** | -0.245*** | -0.245*** | 0.044*** | 0.023*** | 0.024*** | 0.217*** | 0.228*** | 0.226*** |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.004) | (0.004) | (0.004) |
| Gender(Unknown) | -0.294*** | -0.264*** | -0.263*** | 0.034*** | 0.020*** | 0.020*** | 0.181*** | 0.163*** | 0.162*** |
| | (0.005) | (0.005) | (0.005) | (0.004) | (0.004) | (0.004) | (0.005) | (0.005) | (0.005) |
| Log no.pub | -0.226*** | -0.233*** | -0.233*** | -0.129*** | -0.138*** | -0.139*** | 0.495*** | 0.517*** | 0.517*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Proficient | 0.621*** | 0.692*** | 0.303*** | -0.192*** | -0.087*** | 0.000 | -0.257*** | -0.359*** | -0.171*** |
| | (0.004) | (0.005) | (0.009) | (0.004) | (0.004) | (0.008) | (0.005) | (0.005) | (0.010) |
| Developing | 0.531*** | 0.772*** | 0.609*** | 0.097*** | 0.048*** | -0.020+ | -0.603*** | -0.599*** | -0.329*** |
| | (0.009) | (0.011) | (0.014) | (0.008) | (0.009) | (0.012) | (0.011) | (0.012) | (0.017) |
| Lagging | 0.234*** | 0.391*** | 0.306*** | 0.166*** | 0.148*** | 0.020+ | -0.418*** | -0.472*** | -0.163*** |
| | (0.008) | (0.010) | (0.013) | (0.007) | (0.008) | (0.010) | (0.010) | (0.011) | (0.014) |
| Funded | | 1.040*** | 0.893*** | | -0.739*** | -0.743*** | | 0.284*** | 0.408*** |
| | | (0.004) | (0.005) | | (0.003) | (0.004) | | (0.004) | (0.005) |
| Proficient × Funded | | | 0.563*** | | | -0.111*** | | | -0.285*** |
| | | | (0.011) | | | (0.009) | | | (0.012) |
| Developing × Funded | | | 0.290*** | | | 0.168*** | | | -0.529*** |
| | | | (0.021) | | | (0.017) | | | (0.023) |
| Lagging × Funded | | | 0.070*** | | | 0.292*** | | | -0.596*** |
| | | | (0.018) | | | (0.015) | | | (0.020) |
| Num.Obs. | 6540474 | 6583198 | 6583198 | 6457314 | 6521140 | 6521140 | 6540474 | 6583198 | 6583198 |
| R2 | 0.087 | 0.101 | 0.101 | 0.096 | 0.100 | 0.100 | 0.135 | 0.139 | 0.139 |
| R2 Adj. | -0.271 | -0.256 | -0.255 | -0.148 | -0.144 | -0.144 | -0.223 | -0.218 | -0.217 |

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

c

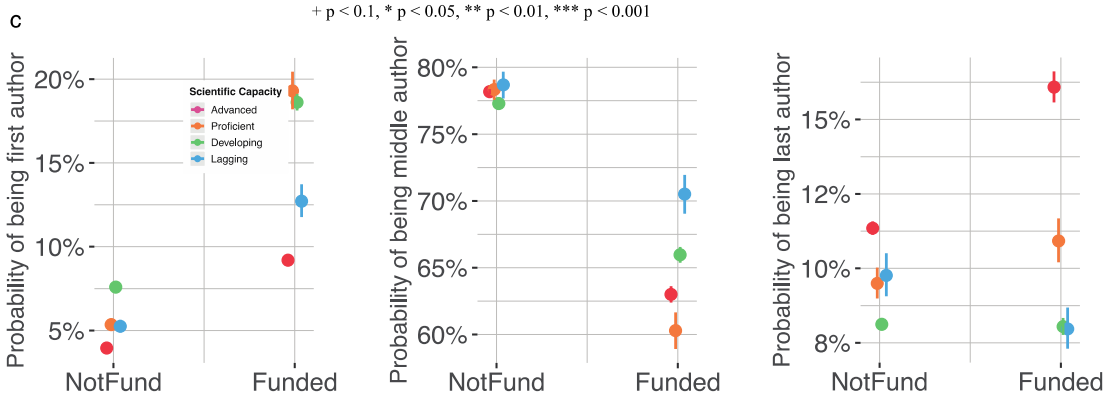
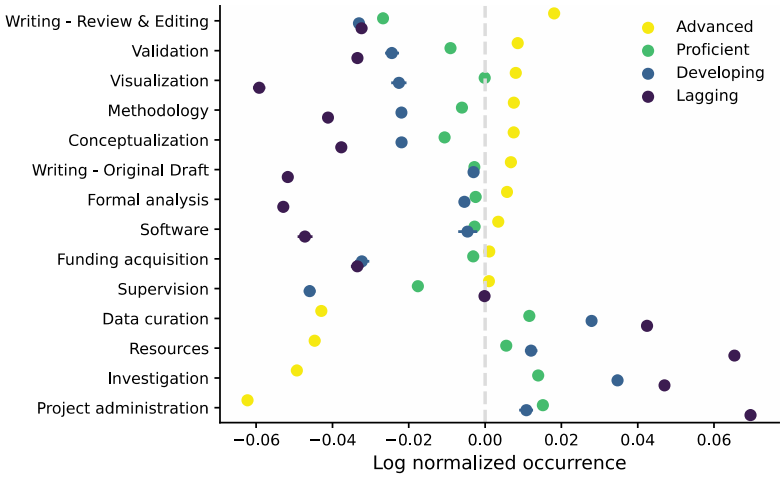



*Figure 3* ***Researchers from scientifically non-advanced countries are less likely to assume the role of the last author, even when they provide funding for the research****. (a) Results of the paper-level fixed-effect regression model. (b)The marginal effect of funding on playing the role of the first author, middle author, and the last author.*

However, due to the lack of large-scale data on detailed author contributions and a clear theoretical understanding of the pathways linking specific contributions to authorship positions, the above regression results may be confounded by the authors' contributions. In other words, the observed effect of nationality on authorship positions reflects both the indirect effect of nationality on authorship via contributions and the direct effect of nationality on authorship positions. To better identify the impact of nationality on both contributions and authorship

positions, we replicate the analysis using a dataset that includes author contributions from PLOS journals (see Data and Methods). The normalized contribution distribution again reveals a clear divide between researchers from scientifically advanced countries and those from less advanced countries. Contributions that researchers from advanced countries tend to make are often underrepresented among researchers from non-advanced countries. Contributions that are underrepresented by researchers from advanced countries are overrepresented by those from non-advanced countries (see Fig. 4a). The paper-level fixed effects regression model further confirms that, after controlling for other confounding factors such as the gender and scientific capacity of authors, researchers from scientifically developing countries and lagging countries are less likely to be associated with intellectual contributions (see Data and Methods, Fig. 4b and SI). Even after accounting for contributions, the regression results still indicate a significant impact of nationality on authorship positions. Researchers from scientifically lagging countries consistently face disadvantages in obtaining both first and last author positions, but are more likely to middle middle author roles (see Fig. 4b).

a




b

| | Intellectual Role | First Author | Middle Author | Last Author |
|---|---|---|---|---|
| Male | 0.051 | -0.422*** | 0.220*** | 0.009 |
| | (0.040) | (0.059) | (0.045) | (0.066) |
| Gender(Unknown) | -0.051 | -0.206* | 0.072 | -0.048 |
| | (0.056) | (0.083) | (0.065) | (0.095) |
| Log no.pub | 0.325*** | -0.362*** | -0.104*** | 0.610*** |
| | (0.013) | (0.016) | (0.012) | (0.022) |
| Proficient | 0.049 | 0.406*** | -0.244*** | 0.077 |
| | (0.061) | (0.071) | (0.064) | (0.079) |
| Developing | -0.458*** | 0.350* | -0.108 | -0.016 |
| | (0.130) | (0.149) | (0.122) | (0.163) |
| Lagging | -0.606*** | -0.350*** | 0.586*** | -0.605*** |
| | (0.078) | (0.094) | (0.078) | (0.109) |
| Leading Role | | 2.197*** | -2.235*** | 1.912*** |
| | | (0.073) | (0.055) | (0.078) |
| Num.Obs. | 25379 | 25278 | 25278 | 25278 |
| R2 | 0.203 | 0.144 | 0.178 | 0.208 |
| R2 Adj. | 0.036 | -0.177 | -0.034 | -0.113 |

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

*Figure 4* ***Researchers from scientifically advanced countries are overrepresented in contributions related to intellectual work.*** *(a) Contributions made by researchers from different scientific groups are normalized by the expected value derived from randomly shuffling the contributions within each paper. A normalized occurrence greater than 0 indicates that researchers from the group are overrepresented in that contribution, while values less than 0 indicate underrepresentation. (b) Paper-level fixed effects regression model.*

Researchers from less-advanced countries not only face disadvantages in assuming the role of the last author, but their research agendas may also be overlooked during international collaboration. By comparing the content similarity of papers published without international collaboration to those produced through such collaboration across countries, the results reveal a

reverse relationship between the similarity of these two sets of papers and the scientific capacity of the countries (see Fig. 4a). In general, domestic publications and international publications share high content similarity in scientifically advanced countries (see Fig. 4a). This similarity decreases as the scientific capacity of country declines, with the lowest similarity observed in scientifically lagging countries (see Fig. 4a). To account for the possibility that these patterns are influenced by the number of publications within each country, we conduct a counterfactual experiment in which each country was assumed to produce the same number of publications in each discipline (see Data and Methods). We then compare the content similarity between national and international publications. The results reveal a consistent pattern: content similarity decreases as the scientific capacity of countries decreases (see SI). However, the average differences within each group narrow to a smaller range. Therefore, we believe that the pattern of content similarity being positively related to a country's scientific capacity is valid, but the actual effect likely falls between the differences observed in the empirical data and those observed in the counterfactual scenario.

Additionally, authorship order influences the content studied in international collaboration. Internationally coauthored publications where authors play the leading role (either as first author or last author) show higher similarity with domestic publications from the same country, compared to publications where authors from that country assume middle authorship roles in the international collaboration (see Fig. 4b). This pattern remains consistent even when we control for the number of publications produced by each country in each discipline (see Data and Methods, SI).

The varying levels of dissimilarity between international and national publications across different scientific capacity groups are also related to countries' specialization in different research areas. As shown in Fig. 4c, as a country's scientific capacity decreases, the content of their research diverges more significantly from that of countries with higher scientific capacity. Generally, scientifically lagging countries exhibit the lowest content similarity with other countries, while scientifically advanced countries show the highest similarity. Given the low topical similarity in lagging countries and the marginalization faced by their researchers in international collaborations, it is not surprising that international coauthored publications from these countries diverge more from their national publications (see Fig. 4c).

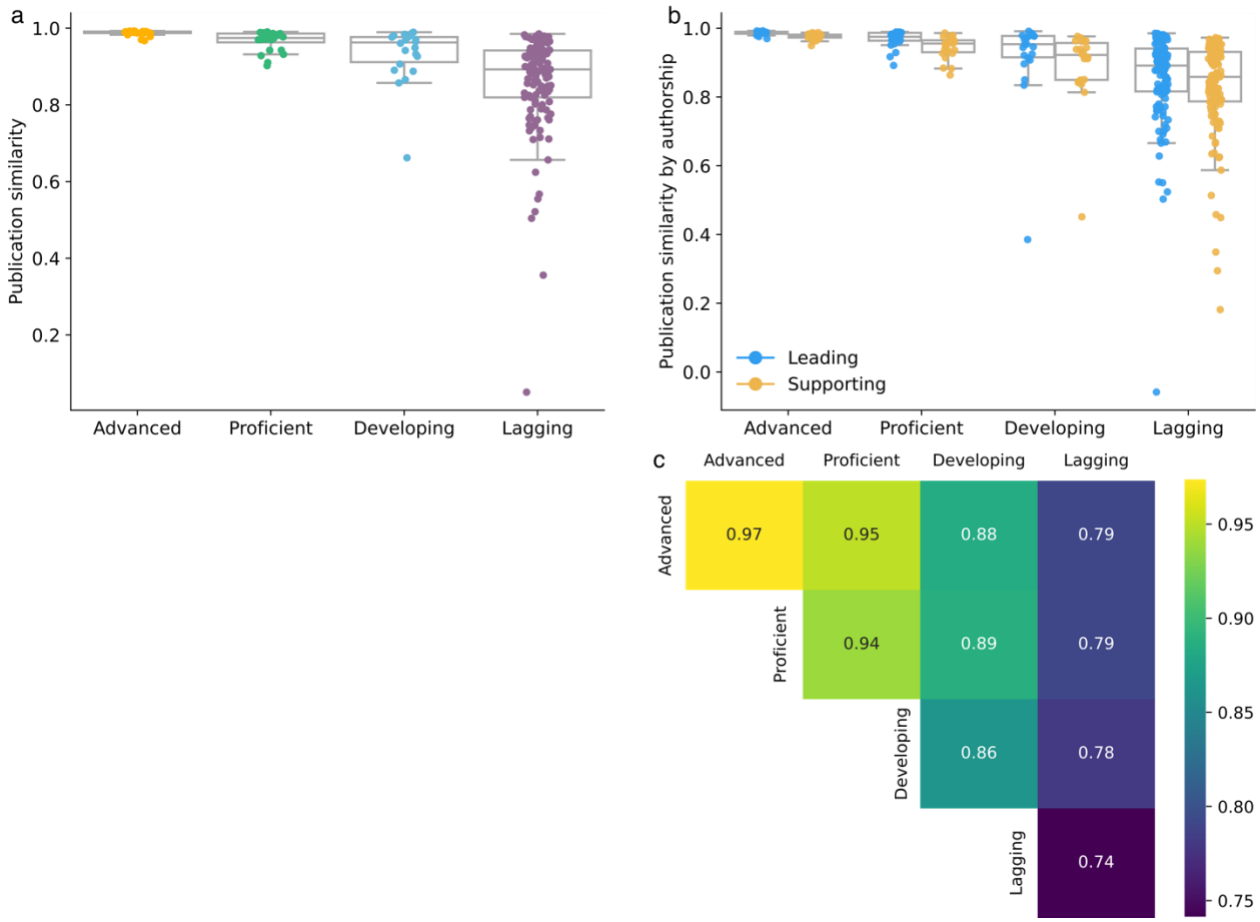



*Figure 5 The similarity between internationally coauthored publications and domestic publications decreases as the scientific capacity of the country decreases. (a) The similarity between internationally coauthored publications and domestic publications within each country, grouped by the country's scientific capacity level. (b) The same similarity while distinguishing whether*

*authors from the focal country play leading roles or supporting roles. (c) Topical similarity of non-internationally collaborated publications between country pairs, aggregated by scientific capacity level.*

## Discussion

Researchers from less-developed countries have constantly raised concerns within the global scientific community regarding the issues such as authorship attribution and the underlying power structure in international collaborations. However, a comprehensive empirical analysis of discrimination against authors from less-developed countries in international collaborations has been lacking. By analyzing publication across all disciplines over 200 countries, we provide a large-scale analysis of the relationship between authors' nationalities and their positions in the collaborative works, along with the consequences. Our findings reveal that authorship order in internationally collaborated papers follows a hierarchical structure, wherein researchers from scientifically proficient and developing countries are more likely to assume the role of first author, researchers from scientifically lagging countries tend to play the role of middle author, and researchers from scientifically advanced countries are more likely to assume the role of last author. Furthermore, even after controlling for influential factors such as authors' scientific capacity and funding status, these patterns remain consistent, reaffirming the existence of a power structure in international collaborations.

Given that authorship is critical for accumulating scientific capital and reputation, and that first and last authors typically receive higher credits for the research[28], our results indicate that researchers from scientifically lagging countries are disadvantaged in obtaining recognition from international collaborations. Moreover, while uneven funding distribution is often blamed for power asymmetry in international collaborations, our results reveal that funding indeed plays a role but is not the sole factors. Unfortunately, funded researchers from lagging countries still

have a lower probability of assuming the role of last author compared to unfunded researchers from advanced countries.

There are several limitations in our analysis. First and foremost, there are substantial differences in authorship attribution practices across disciplines and countries. Disciplines like arts, humanities and social sciences adhere to a more classical notion of authorship, where writing is the primary contribution garnering authorship and other types of work are often unrewarded[24,29]. In contrast, in biology, the first author is typically attributed to the person who performed the majority of the experiments, and the last author is attributed to the head of the laboratory[24]. Although the authorship practice largely depends on the specific norms of a scientific discipline, the pattern of dominant authors (first and last authors) being more likely associated with conceptual tasks and middle authors being associated with technical tasks hold across the entire spectrum of subfields, even in the social science[24]. Therefore, varying authorship practice does not diminish the validity of our findings that researchers from lagging countries are relegated to the supportive role and tasks in collaborations. Meanwhile, another issue with using authorship to infer the collaboration dynamic is that authorship may not always accurately reflect the true contributions of authors, particularly in the cases of 'ghost authorship' and 'guest authorship'. Individual who contributes significantly to the research may be excluded from authorship, while those included as authors may not have made substantial contributions. Although these unethical practices occur, empirical studies indicate that they are not prevalent: analyses of a sample of medical publications found around 19% publications had evidence of guest authors, and around 11% of publications had evidence of ghost authors[30,31]. Given the relative low proportion of

papers with unethical authorship practices, we believe our results are not influenced by such occurrences.

Despite these limitations, our empirical results provide a useful perspective to understand the power asymmetry in current international cooperative frameworks and its adverse effects. The revealed hierarchical structure in scientific collaboration is a reflection of the colonial paradigm evident in broader international cooperative frameworks. Similar power hegemony can be found across a wide range of conditions, such as setting international trade rules, cooperation and negations on crucial issues like climate change, and strict enforcement of intellectual property rights. This prevalent hegemony has historically maintained the dominance of power countries, often at the expense of the sovereignty and self-determination of weaker countries. We hope our study can provide empirical evidence of this neo-colonialism and serve as a starting point for the global community to develop a more equitable and inclusive framework for cooperation.

**Data and Methods**

**Data** The dataset is drawn from the Clarivate Analytics' Web of Science database hosted and managed by the Observatoire des Sciences et des Technologies at the University of Montreal. The Web of Science database contains three main citation indices: The Science Citation Index Expanded, the Social Science Citation Index, and the Arts and Humanities Citation Index. We use all indexed papers, including journal articles and review articles, that were published between 2008 to 2020 which in total contains 19,865,673 papers across 201 countries. After excluding publications with missing information in authorship details, author affiliations and demographic information, the final dataset consists of 17,380,209 publications, which accounted for 86% of the total publications.

For authors with multiple affiliations, we select the affiliation with the highest rank to determine the nationality of the author. Demographic information of authors included in our analysis contains the gender of authors, the total number of publications indexed in WoS database up to 2022, and the first year the author had a publication in WoS database. Authors are disambiguated using the algorithm developed by Caron and van Eck[32]. The gender of author is inferred through their first name[33].

As we utilize the authorship order to infer the collaborative dynamics in international collaboration, publications ordered alphabetically are removed from the analysis. Among the 3,661,174 internationally coauthored publications, 465,185 papers (13% of internationally coauthored publications) have authors ordered alphabetically. For the analysis involving authorship order, only non-alphabetically ordered publications are included. Although it is possible that the last names of authors may have been accidentally ordered alphabetically in the byline, we adhere to the strict criteria to maintain a clean dataset, ensuring that any publications ordered alphabetically were excluded from the analysis.

Information on the funding of a paper was retrieved from the 'Funding Agency' and 'Grant Number' fields in the WoS. We rely on a previously curated dataset containing the country location information of funding agencies to determine the funding source of a publication[34]. Since funding information is retrieved at the paper level, it is infeasible to link the source of funding with specific authors. Therefore, to estimate whether an author provided funding to the paper, we apply a more rough approach by examining whether the funding agency is located in

the same country as the author. Specifically, if at least one funding agency in the paper is locate at the same country as the author, we classify the author as funded.

Countries are classified based on their scientific capacity as developed by Wagner et al[35]. In this classification system, 22 countries are classified as scientifically advanced countries, 24 countries are classified as scientifically proficient countries, 24 countries are classified as scientifically developing countries in the original classification. However, because Yugoslavia and Hong Kong are no longer existing entities, 22 countries are classified as scientifically development countries in our analysis. And the remaining 133 countries are classified as scientifically lagging countries.

To ensure the observed patterns are robust against the classification scheme, we also classify countries based on their income level according to the annual data published by the World Bank. Since our analysis spans over 10 years, countries are assigned to the income level group they most frequently belong to during this period. Consequently, 66 countries are classified as High-income countries, 52 countries are classified as Upper-Middle income countries, 51 countries are classified as Lower-Middle income countries and 32 countries are classified as Low-income countries.

To better measure the causal impact of nationality, author contributions, and authorship positions, we replicate our analysis using a dataset from PLOS journals covering the 2017-2018 period (*N=30770* papers)[26]. This dataset includes detailed author contributions described by the Contributor Roles Taxonomy (CRediT). Since our focus is on the division of scientific labor in international collaborations between researchers from countries with different scientific capacities, we excluded publications with incomplete author demographic information, as well as

national publications and international publications coauthored by countries with the same scientific capacity. This resulted in a final dataset of 4,200 papers.

**Authors distribution across scientific capacity groups.** To calculate the number and proportion of authors involved in international collaborations within each scientific capacity group, we assign authors to individual countries based on their affiliation information. Authors with multiple affiliations are attributed to the country of their first listed affiliation. If an author has multiple first affiliations located at different countries within the same year, the author is counted in each corresponding country. International collaborated publications are those where authors, after being assigned to specific countries, come from different countries. Authors of these publications are considered to be participating in international collaboration. After assigning authors to specific countries, we aggregated the counts based on the scientific capacity classification of each country. The proportion in each group is calculated by dividing the number of authors in each classification group by the total number of authors engaged in international collaboration for that year.

**Authorship occurrence normalization.** To calculate the expected value of a specific authorship within countries, we perform 20 random shuffling of the authorship order within each internationally coauthored publications. The normalized authorship occurrence is then derived using the formula:

$$T_{c,i} = \frac{W_{c,i}}{E_{c,i}}$$

Where $W_{c,i}$ denotes the frequency of researchers from country $c$ appearing in the authorship $i$, $E_{c,i}$ represents the corresponding value obtained from random shuffling, and $T_{c,i}$ is the normalized occurrence of researchers from country $c$ in authorship $i$. A $T_{c,i}$ value large than 1

suggests researchers appear in the authorship more frequently than expected, whereas a $T_{c,i}$ value bellow 1 implies researchers from country *c* are less frequent in authorship *i.*

We applied a similar normalization to the PLOS dataset by randomly shuffling the CRediT contributions of authors within each paper, while keeping the number of contributions made by each author unchanged and ensuring that authors were not assigned duplicate contributions.

**Paper-level fixed-effect regression model.** In an ideal scenario, authorship order should only reflect the contributions made by authors[36]. However, in reality, determining authorship order is a complex process influenced by various factors. Empirical studies have revealed the relationships between authorship order and factors such as gender[27], age[37] and professional rank[37]. Therefore, to better reveal the causal relationship between potential influential factors and the authorship order in international coauthored publications, we employ a fixed-effect regression model, treating each individual paper as the analysis unit. Drawing upon these empirical findings, our conceptual model posits that authorship order is determined by authors' contributions, their gender, nationality, and scientific capacity (see Fig. 3a). We use the number of publications produced by each author up to 2022 as a proxy for their scientific capacity. However, due to the lack of fine-grained data on authors' contributions for each internationally collaborated paper, our findings regarding the causal impact of nationality on authorship position reflect a combination of the indirect impact via contributions and the direct impact on authorship order.

Specifically, the paper-level fixed-effect regression model defines playing a specific authorship role as the binary outcome while adjusting for paper-level heterogeneity according to the following model:

$$y_i = \log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \cdots + \alpha_i + \varepsilon_i$$

Where $p$ is the probability that an author is playing the role of a specific authorship, $x_{i,j}$ represents independent variables, $\alpha_i$ accounts for paper-fixed effect, and $\varepsilon_i$ is the error term.

To better account for the influence of contributions on authorship position, we apply the same model to the PLOS dataset, considering the specific contributions made by each author. Since authorship order is not determined by a single contribution but is generally associated with the intellectual contributions of the authors, we group the contributions into intellectual-related and non-intellectual-related categories. Building on existing studies[38,39], we classify four types of contributions—Writing-Review & Editing, Methodology, Conceptualization, and Supervision—as intellectual-related tasks, while the remaining tasks are categorized as non-intellectual-related. We then use whether an author performed intellectual-related tasks as a dependent variable to examine the impact of nationality on the division of scientific labor. Given that an author can contribute to multiple tasks, we classify an author as having performed intellectual-related tasks if they contributed to at least one task in the intellectual-related set. This classification is subsequently used as an independent variable in models where authorship order was the dependent variable.

**Measuring similarity between papers.** The content of a publication is represented by a 64-dimensional vector computed from the SPECTER model using title and abstract information[40].

To estimate the similarity between internationally coauthored publications and non-international coauthored publications, we compute the cosine distance between the mean vectors of publications authored internationally and those authored domestically within each discipline for every country. Furthermore, to identify the role of authorship in shaping international collaborations, we distinguish between publications led by authors and those supported by authors. “Led by authors” publications indicate that authors from the focal country played a leading role in the publication, while “supported by authors” publications refer to those in which authors from the focal country played a supportive role. Considering that the first and last authors typically contribute most to the research and often have the greatest influence on the research direction, publications “led by authors” are defined as those where authors from the focal country are either the first or last author, while “supported by authors” publications are those where authors from the focal country are the middle author. If authors from the focal country assume both leading and supporting roles in the publication, we assign the publication to the “led by authors” group.

To ensure that the observed pattern is not an artifact of the number of publications within each country, we conduct a counterfactual experiment assuming that each country produces the same number of publications in each discipline. First, we exclude disciplines from a country where the number of publications is fewer than 20. For the remaining disciplines, we then sample 100 papers with replacement from the raw publication list. Finally, we compare the content similarity using these sampled 100 papers.

**References**

1. Stefanoudis, P. V. *et al.* Turning the tide of parachute science. *Curr. Biol.* **31**, R184–R185 (2021).

2. Polonsky, M. J., Garma, R. & Mittelstaedt, J. D. An examination of the globalisation of authorship in publishing in 20 leading marketing journals. *Eur. Bus. Rev.* **18**, 437–456 (2006).
3. García-Carpintero, E., Granadino, B. & Plaza, L. M. The representation of nationalities on the editorial boards of international journals and the promotion of the scientific output of the same countries. *Scientometrics* **84**, 799–811 (2010).
4. Gomez, C. J., Herman, A. C. & Parigi, P. Leading countries in global science increasingly receive more citations than other countries doing similar research. *Nat. Hum. Behav.* **6**, 919–929 (2022).
5. Tuyishime, H. *et al.* Authorship Distribution and Under-Representation of Sub-Saharan African Authors in Global Oncology Publications. *JCO Glob. Oncol.* e2200020 (2022) doi:10.1200/GO.22.00020.
6. Garbern, S. C. *et al.* Authorship representation in global emergency medicine: a bibliometric analysis from 2016 to 2020. *BMJ Glob. Health* **7**, e009538 (2022).
7. Dahdouh-Guebas, F., Ahimbisibwe, J., Van Moll, R. & Koedam, N. Neo-colonial science by the most industrialised upon the least developed countries in peer-reviewed publishing. *Scientometrics* **56**, 329–343 (2003).
8. Marks, R. A., Hotaling, S., Frandsen, P. B. & VanBuren, R. Representation and participation across 20 years of plant genome sequencing. *Nat. Plants* **7**, 1571–1578 (2021).
9. Gewin, V. Four global-south researchers making cross-border collaborations count. *Nature* **624**, S2–S6 (2023).
10. Bendana, C. I advocate an African research agenda for African development. *Nature* (2023) doi:10.1038/d41586-023-03241-w.

11. Wallerstein, I. *The Modern World-System I: Capitalist Agriculture and the Origins of the European World-Economy in the Sixteenth Century*. (University of California Press, 2011).
12. Chinchilla-Rodríguez, Z., Sugimoto, C. R. & Larivière, V. Follow the leader: On the relationship between leadership and scholarly impact in international collaborations. *PLOS ONE* **14**, e0218309 (2019).
13. Marginson, S. What drives global science? The four competing narratives. *Stud. High. Educ.* **0**, 1–19 (2021).
14. Hwang, K. International Collaboration in Multilayered Center-Periphery in the Globalization of Science and Technology. *Sci. Technol. Hum. Values* **33**, 101–133 (2008).
15. Ramírez-Castañeda, V. *et al.* A set of principles and practical suggestions for equitable fieldwork in biology. *Proc. Natl. Acad. Sci.* **119**, e2122667119 (2022).
16. Basalla, G. The Spread of Western Science. *Science* **156**, 611–622 (1967).
17. Crane, J. T. Unequal 'Partners'. AIDS, Academia, and the Rise of Global Health. *Behemoth* **3**, (2010).
18. Raman, S. What it means to practise values-based research. *Nature* (2023) doi:10.1038/d41586-023-01878-1.
19. Gewin, V. Invest the time to build trust among marginalized research participants. *Nature* **612**, 177–178 (2022).
20. Gewin, V. Decolonization should extend to collaborations, authorship and co-creation of knowledge. *Nature* **612**, 178–178 (2022).
21. Schubert, T. & Sooryamoorthy, R. Can the centre–periphery model explain patterns of international scientific collaboration among threshold and industrialised countries? The case of South Africa and Germany. *Scientometrics* **83**, 181–203 (2009).

22. Choung, J.-Y. & Hwang, H.-R. National Systems of Innovation: Institutional Linkages and Performances in the Case of Korea and Taiwan. *Scientometrics* **48**, 413–426 (2000).
23. Choung, J.-Y. & Hwang, H.-R. The evolutionary patterns of knowledge production in Korea. *Scientometrics* **94**, 629–650 (2013).
24. Larivière, V. *et al.* Contributorship and division of labor in knowledge production. *Soc. Stud. Sci.* **46**, 417–435 (2016).
25. Tarkang, E. E., Kweku, M. & Zotor, F. B. Publication Practices and Responsible Authorship: A Review Article. *J. Public Health Afr.* **8**, 723 (2017).
26. Larivière, V., Pontille, D. & Sugimoto, C. R. Investigating the division of scientific labor using the Contributor Roles Taxonomy (CRediT). *Quant. Sci. Stud.* **2**, 111–128 (2021).
27. The gendered nature of authorship | Science Advances. https://www.science.org/doi/full/10.1126/sciadv.abe4639.
28. Rees, C. A. *et al.* Importance of authorship and inappropriate authorship assignment in paediatric research in low- and middle-income countries. *Trop. Med. Int. Health* **24**, 1229–1242 (2019).
29. Sugimoto, C. R. & Vincent, L. *Measuring Research: What Everyone Needs to Know*. (Oxford University Press, 2018).
30. Flanagin, A. *et al.* Prevalence of Articles With Honorary Authors and Ghost Authors in Peer-Reviewed Medical Journals. *JAMA* **280**, 222–224 (1998).
31. Wislar, J. S., Flanagin, A., Fontanarosa, P. B. & DeAngelis, C. D. Honorary and ghost authorship in high impact biomedical journals: a cross sectional survey. *BMJ* **343**, d6128 (2011).

32. Caron, E. & Eck, N. J. van. Large scale author name disambiguation using rule-based scoring and clustering. in *Proceedings of the 19th international conference on science and technology indicators* (Leiden, 2014).
33. Larivière, V., Ni, C., Gingras, Y., Cronin, B. & Sugimoto, C. R. Bibliometrics: Global gender disparities in science. *Nature* **504**, 211–213 (2013).
34. Miao, L., Larivière, V., Wang, F., Ahn, Y.-Y. & Sugimoto, C. R. Cooperation and interdependence in global science funding. Preprint at https://doi.org/10.48550/arXiv.2308.08630 (2024).
35. Caroline S. Wagner, Irene Brahmakulam, Brain Jackson, Anny Wong, & Tatsuro Yoda. *Science and Technology Collaboration: Building Capacity in Developing Countries?* (2001).
36. Helgesson, G. & Eriksson, S. Authorship order. *Learn. Publ.* **32**, 106–112 (2019).
37. Costas, R. & Bordons, M. Do age and professional rank influence the order of authorship in scientific publications? Some evidence from a micro-level perspective. *Scientometrics* **88**, 145–161 (2011).
38. Robinson-Garcia, N., Costas, R., Sugimoto, C. R., Larivière, V. & Nane, G. F. Task specialization across research careers. *eLife* **9**, e60586 (2020).
39. Xu, F., Wu, L. & Evans, J. Flat teams drive scientific innovation. *Proc. Natl. Acad. Sci.* **119**, e2200927119 (2022).
40. Cohan, A., Feldman, S., Beltagy, I., Downey, D. & Weld, D. S. SPECTER: Document-level Representation Learning using Citation-informed Transformers. Preprint at http://arxiv.org/abs/2004.07180 (2020).

# Supplementary Information

## Analysis with publications across different groups

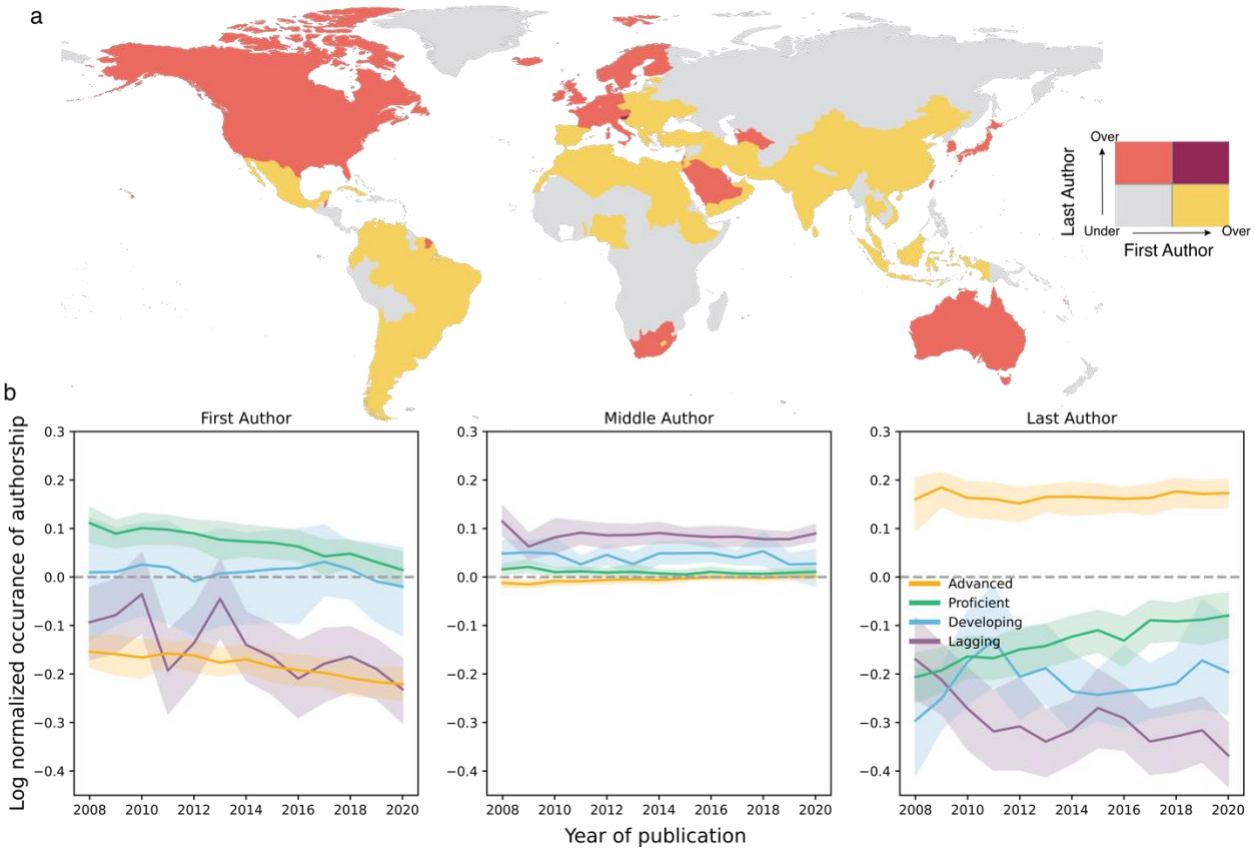



*Figure 6 Authorship representation in publications collaborated by countries with different scientific capacity.*

*Table 1 Regression models with CrediT task as dependent variables.*

| | Investigation | Writing - Original Draft | Writing - Review & Editing | Supervision | Formal analysis | Validation | Methodology |
|---|---|---|---|---|---|---|---|
| Male | -0.324*** | -0.258*** | 0.048 | 0.170*** | -0.128*** | -0.107* | -0.124** |
| | (0.039) | (0.039) | (0.048) | (0.044) | (0.037) | (0.049) | (0.038) |
| Gender(Unknown) | -0.092+ | -0.173** | -0.264*** | -0.114+ | -0.245*** | -0.189** | -0.232*** |
| | (0.054) | (0.056) | (0.070) | (0.063) | (0.052) | (0.070) | (0.055) |
| Log no.pub | -0.211*** | 0.082*** | 0.540*** | 0.720*** | -0.050*** | 0.126*** | 0.067*** |
| | (0.013) | (0.011) | (0.015) | (0.018) | (0.011) | (0.015) | (0.012) |
| Proficient | 0.411*** | -0.003 | -0.579*** | 0.220*** | -0.048 | -0.052 | -0.104+ |
| | (0.056) | (0.052) | (0.068) | (0.058) | (0.055) | (0.069) | (0.056) |
| Developing | 0.652*** | -0.276** | -0.628*** | 0.042 | -0.397*** | -0.370** | -0.603*** |
| | (0.113) | (0.104) | (0.128) | (0.110) | (0.108) | (0.136) | (0.116) |
| Lagging | 0.396*** | -0.790*** | -0.817*** | 0.325*** | -1.069*** | -0.703*** | -0.961*** |
| | (0.074) | (0.069) | (0.093) | (0.077) | (0.070) | (0.089) | (0.075) |
| Num.Obs. | 26366 | 29918 | 20803 | 27753 | 28856 | 18111 | 26500 |
| R2 | 0.201 | 0.161 | 0.299 | 0.271 | 0.163 | 0.181 | 0.183 |
| R2 Adj. | 0.021 | -0.052 | 0.119 | 0.071 | -0.028 | -0.016 | 0.002 |

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

*Table 1 Continued*

| | Funding acquisition | Project administration | Resources | Data curation | Conceptualization | Visualization | Software |
|---|---|---|---|---|---|---|---|
| Male | 0.059 | -0.056 | 0.211*** | -0.247*** | -0.040 | -0.156** | 0.313*** |
| | (0.047) | (0.044) | (0.047) | (0.041) | (0.038) | (0.054) | (0.062) |
| Gender(Unknown) | -0.019 | -0.050 | 0.108+ | -0.071 | -0.048 | -0.324*** | 0.257** |
| | (0.069) | (0.062) | (0.065) | (0.055) | (0.056) | (0.078) | (0.088) |
| Log no.pub | 0.676*** | 0.308*** | 0.319*** | -0.212*** | 0.462*** | -0.039* | -0.145*** |
| | (0.019) | (0.014) | (0.016) | (0.012) | (0.013) | (0.015) | (0.018) |
| Proficient | 0.448*** | 0.564*** | 0.471*** | 0.277*** | 0.113* | -0.010 | -0.043 |
| | (0.063) | (0.061) | (0.069) | (0.060) | (0.055) | (0.072) | (0.088) |
| Developing | 0.120 | 0.391*** | 0.469*** | 0.219+ | -0.170 | -0.429** | -0.232 |
| | (0.132) | (0.108) | (0.138) | (0.117) | (0.110) | (0.147) | (0.167) |
| Lagging | -0.181* | 0.801*** | 0.792*** | 0.100 | -0.498*** | -1.099*** | -0.847*** |
| | (0.087) | (0.075) | (0.097) | (0.078) | (0.071) | (0.108) | (0.116) |
| Num.Obs. | 25000 | 23948 | 20168 | 25091 | 28743 | 16035 | 12489 |
| R2 | 0.236 | 0.181 | 0.197 | 0.186 | 0.216 | 0.163 | 0.136 |
| R2 Adj. | 0.017 | -0.026 | 0.008 | -0.003 | 0.031 | -0.050 | -0.093 |

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

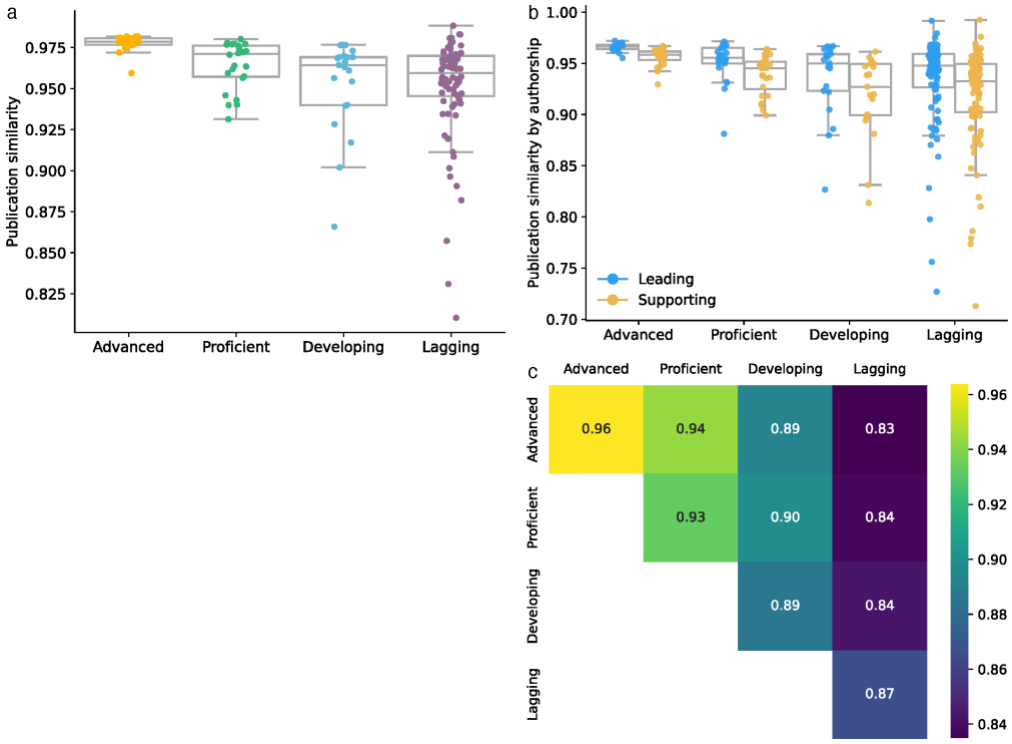



*Figure 7 Content similarity derived from resampled publications.*