Two separated worlds - on the international influence of life science and biomedical research

Zuguang Gu
National Center for Tumor Diseases (NCT)
Heidelberg, Germany
z.gu@dkfz.de

May 16, 2024

Abstract

We have introduced a new metric, "citation enrichment", to measure the country-to-country influence using citation data. This metric evaluates the preference of a country to cite the other country compared to a random citation process. We applied the citation enrichment method on over 14 million publications in the life science and biomedical fields and we have the following findings: 1) The globe is split into two separated worlds where developed western countries exhibit an overall mutual negative influence on the remaining of the world; 2) Within each world, countries form clusters based on their mutual influences and the classification fits perfectly to their geographical and cultural proximity; 3) The two worlds exhibit distinct patterns of the citation balance among countries. We have constructed a comprehensive world map of scientific influence which greatly enhances the deep understanding of the international exchange of scientific knowledge. The citation enrichment metric is developed under a well-defined statistical framework and it can be extended to a general and powerful tool for research in bibliometrics and related fields.

***K*eywords** Citation analysis · Citation enrichment · Citation balance · International influence · Bibliometrics

1 Introduction

Scientific publications serve as the primary medium for storing and disseminating scientific knowledge. They play a vital role in documenting research findings, advancements, and insights across various fields of research. These publications encompass a wide range of formats, including research articles, reviews, conference papers, reports and scientific books. Among them, knowledge flows are mainly represented as citations between publications where researchers use them as an acknowledgement of prior research, utilization of established methods and tools, and validation for their scientific findings. In short, citations can be used as a proxy for measuring the scientific influence. In various evaluation systems, citation has been frequently used as an important indicator for ranking publications, authors, journals or institutions. However, despite its simplicity and popularity, the citation has been criticized for its inappropriateness in measuring the influence of scientific outcomes, including the bias towards established researchers and institutions, preference for journals with high impact factors (Tahamtan and Bornmann 2018; Urlings et al. 2021), or overestimation from self-citations and citation manipulation (Wren and Georgescu 2022). What's more, even in an individual publication, the cited publications do not influence the citing publication equally (Zhang and Wu 2021). Nevertheless, on the macro-level, e.g., when evaluating the scientific influence of a country, citation is still one of the most widely used metrics for studying the patterns and dynamics of scientific influence over the globe.

There are a large number of bibliometrics methods that measure a country's scientific influence from citation data (Radicchi and Castellano 2015). The most used methods are count or fraction-based (Donthu et al. 2021), such as the total number or the fraction of international citations of a country. There are also metrics with more complex forms. For example, for a country denoted as c and all the international publications that are not published in c as D, Hassan and Haddawy (2013) measured the impact of c as the fraction

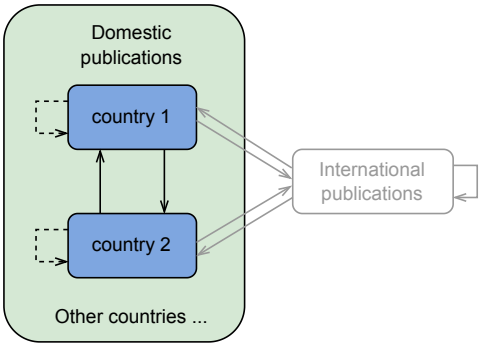


Figure 1: Categories of scientific publications. Domestic publications: papers where authors (or the majority of them) are from the same country; International publications: papers where authors are from multiple countries, i.e., via international collaborations. Arrows represent directions of citations. International publications and their citation relations are colored in grey because they are not included in this study.

of the citations where D cites c to the total citations from D. However, this metric generates tiny values for countries with small citation volumes because the total number of citations from D is usually huge, which makes the interpretation of the impact values difficult. Carammia (2022) measured the country-level international influence in the political science field by the number of papers published in high-impact journals which are believed to be internationally recognized. The limit of this study is, there was no quantitative measurement of the internationization of these journals and the relation of the internationization between journals and papers still needs more careful explorations. Lancho Barrantes et al. (2012) proposed a metric called "impact on papers per paper" by taking the citation network as a bipartite graph where cited papers and citing papers are the two distinct sets[1], then the metric is defined as the number of links divided by the number of cited papers and citing papers from a country, i.e., the average number of citations per cited paper and per citing paper. It is easy to see that if the bipartite graph is fully connected, the metric has a value of 1, but in real cases, the bipartile citation graph is sparse, which makes the metric value very small and the interpretation of it not straightforward.

For a country c, the majority of current studies summarize c's scientific influence by aggregating the corresponding citation data into a single metric (Bogart et al. 2021). Such approaches provide an easy way to compare and rank countries, and to study the longitudinal dynamics of country-level influence (Abramo, D'Angelo, and Carloni 2019; Carammia 2022; Hassan and Haddawy 2013; Lancho Barrantes et al. 2012), but it has obvious limitations. First, the levels of scientific activities are not equally distributed over the world. The majority volumes of scientific knowledge are only contributed by a small number of science-advanced countries. Simply averaging all citations of c without distinguishing country-specific differences of c's citing countries neglects effects from "scientifically small countries". Second, patterns of the scientific knowledge flow may be different in different regions of the world, due to their political, social and cultural diversities. Aggregating citations into a single value makes it impossible to study the interesting region-specific science dissemination patterns from c's citing countries. To reveal such details missing in current studies, it is necessary to extend the country-wise analysis into two dimensions where a country-to-country influence analysis is conducted. Unfortunately, there are very few studies focused on this direction. Abramo, D'Angelo, and Carloni (2019) measured the balance of knowledge flow of a country as the difference between its international citations and international references. Although the focus of that study was still to assign a general "balance value" country-wise, it can be extended to the country-to-country analysis by only looking at a single second country instead of merging all international countries. Nevertheless, there were only four countries investigated in the study which is far from being enough for discovering general patterns on the global level.

All publications can be categorized into two groups: domestic publications where authors are mainly from the same country, and international publications where the studies are conducted by authors from multiple countries (Figure 1). In other words, domestic publications correspond to scientific knowledge generated within individual countries while international publications correspond to the knowledge shared between countries. Citations are directional and link publications within and between both groups. In international

[1]A publication can exist in both sets but marked as citing and cited publications.

publications, it is hard to associate them with a single country, thus the source country of the scientific knowledge is ambiguous. On the other hand, globally, domestic publications are the major type of publications (Chinchilla-Rodríguez 2019; Lancho Barrantes et al. 2012). Thus, in this study, we analyzed the global scientific influence only with domestic publications and the citation flows among them (green box in Figure 1).

The most important part of constructing a worldwide scientific influence map is to let all countries be measured under the same scale or normalized to the same baseline, so that further data analysis approaches such as clustering can be applied to them to reveal more specific patterns of countries. The count-based metrics such as the total number of international citations are heavily affected by a country's absolute citation volume where countries such as the US are always found on the top of the most influential countries (Carammia 2022). Fraction-based metrics such as the fraction of international citations are relatively more robust, but they still have a negative correlation to the citation volumes (e.g., in Figure 3C-D). To solve these problems, in this study, we proposed a new country-to-country metric named "*citation enrichment*" which quantitatively measures the preference of a country citing or cited by another country. Being different from other metrics used in current studies, the citation enrichment metric aims to evaluate the influence from a new aspect by answering the question of "is a country *preferable* to cite another country" by comparing it to a random citation process. Methodologically, it has two advantages. First, it is a relative metric of which all countries are against the same baseline where publications cite randomly. Second, the random model considers the citation volumes of countries and the metric corrects for the effects from larger countries that have higher probabilities of being cited. The metric can measure both positive and negative influences where a positive influence corresponds to more citations than random and vice versa. The enrichment metric has a well-defined form for theoretical statistical analysis and it has already been widely applied in other fields such as biological data analysis (Rivals et al. 2006).

We applied the citation enrichment analysis on more than 14 million publications from 2000 to 2023 in the life science and biomedical fields. The aim of this empirical study is not to rank individual countries, but to provide a comprehensive world map of international scientific influences. We illustrated the mutual influences of 72 countries covering six continents. Using the citation enrichment metric, we discovered a new and interesting pattern of worldwide knowledge flow where the globe is split into two separated worlds. The developed western world is isolated from the remaining of the world where the two worlds show overall negative influence on each other, i.e., they have weaker scientific interactions than random. We discovered that within the two separated worlds, countries can be further clustered into groups where countries within groups have stronger mutual positive influences. Very interestingly, the clustering of countries agrees perfectly with their geographical closeness and cultural similarity. This agreement implies that the preference of the citation process among countries may have a strong association with their cultural identifications. Last, we further studied the balance of mutual influence of countries, i.e., the balance of a country's in- and out-flows of scientific knowledge, which is rarely discussed in current studies. We found the two separated worlds have distinct patterns of balancing the knowledge flows in individual countries, reflecting the dynamics of the scientific influence expansion in the two different world systems.

It is worth mentioning that there are also a large number of current studies focused on international collaborations (Vieira 2023; Lancho Barrantes et al. 2012) where the authors of a publication are from multiple countries (white box in Figure 1). Although international collaborations are not included in this study, it can be expected that there are very close associations between international influences and collaborations. They can be thought of as two forms of scientific knowledge exchange. Nation-level cooperations promote the dissemination of science in a direct way where researchers from different countries directly work together on the same scientific topic and share their intelligence, but also in an indirect way in the forms of lectures, workshops and conferences where researchers know each other's work more quickly, which eventually promotes more mutual citations in publications. Nevertheless, in Section 5, we will discuss the possibility of applying the enrichment method to reveal the patterns of worldwide country-to-country collaborations.

The citation enrichment analysis presented in this study, being different from traditional bibliometrics approaches, provides a new aspect of evaluating scientific influence among countries. We demonstrated the advantages of citation enrichment analysis in that it reveals meaningful and well-structured patterns of global scientific influences that are missing or have not been well-explored in current studies. The well-established statistical framework we have proposed in this study can be extended to a general tool for studying scientific influence on various levels. We believe the citation enrichment analysis will be a powerful tool for bibliometrics studies to generate interesting results in related fields.

2 Data and Preprocessing

2.1 Data source

PubMed is the world's largest public database of literature for biomedical and life science research. We downloaded XML files from the PubMed FTP[2] with file indices from 1 to 1366 (by 2024-04-07). Publications between the years 2000 and 2023, written in English, and with numbers of authors between 1 and 50 were included. 21,366,669 publications were used as the starting dataset of this study.

2.2 Country name identification

The identification of country names from affiliation texts is dictionary-based. In a large number of bibliometrics studies, countries were directly matched from the affiliation texts by a list of pre-compiled country names, e.g., in the tool *bioliometrix* (Aria and Cuccurullo 2017). This approach requires affiliations to be well-formatted. It can almost be ensured in recent publications in which journals have more strict formatting requirements for electronic archiving, but affiliations are more loosely formatted in earlier publications. Country names may have aliases, e.g., United States of America/United States/USA/US or Germany/Deutschland, thus, an incomplete list of country names may cause countries to fail to be identified. It also occurs often that country names are not even encoded in affiliations. For example, a large number of publications from the US only encode affiliations at the state level. Therefore, we built complex dictionaries that contain information on various levels for constructing more comprehensive mappings between affiliations and countries, which include names of countries[3], cities[4], institutions, states/provinces[5], and email domain names. We also manually collected 689 mappings on various levels to official country names.

98.1% of all affiliations were mapped to single countries. We manually validated the identified countries by randomly selecting 1,000 affiliation texts. Since all 1,000 affiliations were correctly identified, we estimate the misidentification rate to be less than 0.1%.

2.3 Domestic and international publications

After the countries were identified from authors' affiliations, in a publication, each author was associated with no country (if the author has no affiliation information or the country can not be identified), a single country or multiple countries. We next identified whether the publications are domestic (i.e., all or most of the authors are from the same country), or international (i.e., authors are from multiple countries). The identification of country types was based on the following rules, applied sequentially:

1. If the country information for all authors is missing, the publication is marked as "empty".
2. If more than one different countries are associated with the first author, the publication is marked as "international". Note the country information for the first author is allowed to be missing. If it is missing, we go to the next rule.
3. If the number of countries associated with the last author is not one, the publication is marked as "international". We assume the last author is the most senior in the publication and should be associated with exactly one country if the publication is domestic.
4. If the country of the first author is different from the last author, the publication is marked as "international".
5. If there are more than two authors, the publication is associated with a country as a domestic publication only when all or the majority of the authors are from the same country. To evaluate it, we first associated a single country with each author in the publication. For an author in the middle of the author list with multiple countries, it was associated with the same country as the last author if they share the same one, or else a random country was selected for this author. Then, the country (apparently it is the country of the last author) is assigned to the publication only if no less than 80% of authors (i.e., at least four out of five authors) are associated with it, or else the publication is marked as "international".
6. If a publication does not satisfy rules 1-5, it is marked as "domestic".

[2] `https://ftp.ncbi.nlm.nih.gov/pubmed/baseline/`, `https://ftp.ncbi.nlm.nih.gov/pubmed/updatefiles/`.
[3] `https://worldpopulationreview.com/country-rankings/official-names-of-countries`.
[4] `https://github.com/w8r/k-d-tree/blob/master/data/data.tsv`.
[5] `https://github.com/rinvex/universities/`.

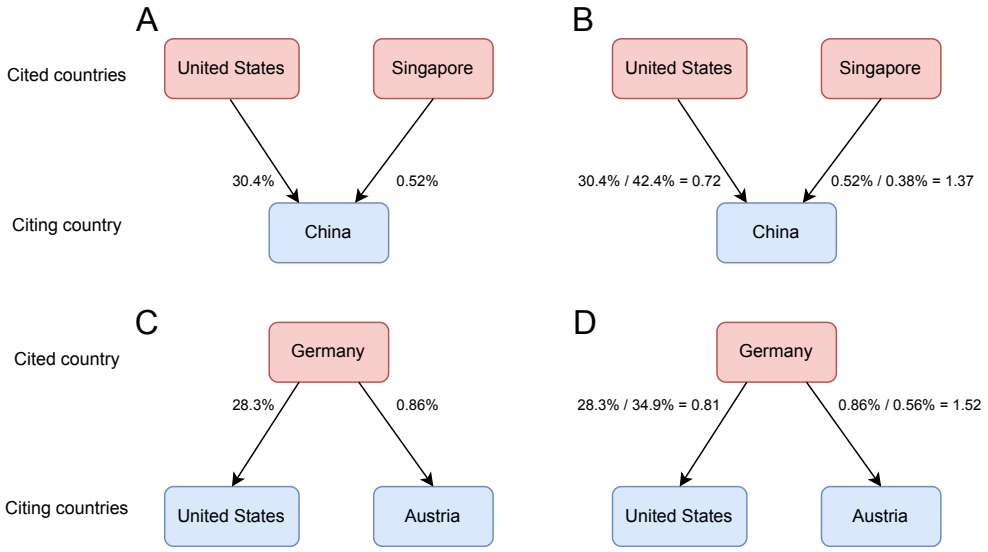


Figure 2: The motivation of using fold enrichment to measure the scientific influence from citation data. A) 30.4% and 0.52% of the citations made by China reference the US and Singapore. B) The fold enrichment of China citing the US and Singapore are 0.72 and 1.37. C) 28.3% and 0.86% of Germany's citations are made by the US and Austria. D) The fold enrichment of Germany being cited by the US and Austria are 0.81 and 1.52.

Additionally, in this study, we restricted our analysis to the publications only published in the World of Science (WoS) core journals collection[6]. We also removed isolated publications that do not cite or are not cited by any other publication in the dataset. Finally, there were 139,760,456 citations from 14,184,112 publications and each publication was associated with a single country.

2.4 Data availability

The processed dataset is available at `https://zenodo.org/doi/10.5281/zenodo.10981961`. The scripts for reproducing the complete analysis in this study are available at `https://github.com/jokergoo/citation_analysis`.

3 Methods

3.1 Motivation

Assume publications from country A cite a total of n times where they cite publications from country B for k times, then it is natural to use the fraction of A citing B to the total number of citations from A, i.e., k/n, as the metric of how B influences A. As an example in Figure 2A, for all citations made by China, 30.4% of them reference the US, which is the largest source of citing countries of China, even larger than China citing itself (25.6%). As a comparison, only 0.52% of citations from China reference Singapore. It seems reasonable to conclude that the US has a dominantly stronger influence than Singapore on China by directly comparing the magnitude of the citing fractions. However, notice that the total volumes of global citations to the US and Singapore also differ significantly (59.2M vs. 0.53M), which results in, 42.4% of the global citations referencing the US and only 0.38% of them referencing Singapore. Let's consider a random scenario where a publication randomly picks another publication to cite, and then naturally countries with larger citation volumes have higher probabilities of being cited. Country-wise, taking the US as an example, the global citing fraction can be used as an estimation of the probability ($p = 0.424$) of the US being cited in the random citation model. Then, if there is no preference for citing the US or other countries, all the countries should have an equal probability of $p = 0.424$ of citing the US. Thus, we can measure the preference of China citing the US as the fold enrichment of the observed citing fraction to the expected citing fraction,

[6] `https://mjl.clarivate.com/collection-list-downloads`.

i.e., 30.4%/42.4% = 0.72 (Figure 2B). The fold enrichment is less than one, and under this context, we can conclude China is negatively influenced by the US where it cites the US less frequently than random. As a comparison, the fold enrichment of China citing Singapore is 0.52%/0.38% = 1.37 which is larger than one, meaning China is positively influenced by Singapore. Overall, we can conclude China is more tightly influenced by Singapore than the US because the former is positive and the latter is negative.

It is similar if we look at the citation enrichment of how a country is cited, i.e., how it distributes its influence to other countries. As an example in Figure 2C, for all citations that Germany receives, 28.3% of them are made by the US, which is the largest country that cites Germany, even larger than Germany being cited by itself (17.6%). However, 34.9% of the global citations are made by the US, which makes the fold enrichment of Germany cited by the US to be 28.3%/34.9% = 0.81, meaning Germany negatively influences the US (Figure 2D). As an opposite example, 0.86% of Germany's citations and 0.56% of global citations are made by Austria, generating a fold enrichment of 0.86%/0.56% = 1.52, implying that Germany has a positive influence on Austria (Figure 2D).

For both examples, the citation enrichment highlights the tight and positive relations between Singapore and China, and also between Germany and Austria, although their absolute citation volumes differ greatly. The results are very meaningful because obviously, the two pairs of countries have great closeness in their economics, culture, and science.

3.2 Citation enrichment

Formally, to calculate the fold enrichment of how country A cites country B or how B is cited by A, let's denote the total number of global citations as N, the total number of citations made from A as n (i.e., the total number of references of A), the total number of citations where B is cited as m, the number of citations of A citing B as k, then the number of citations in different categories can be written into the 2×2 contingency table in Table 1.

	Country B (cited)	Other countries (cited)	Total
Country A (citing)	k	$n-k$	n
Other countries (citing)	$m-k$	$N+k-m-n$	$N-n$
Total	m	$N-m$	N

Table 1: The 2×2 contingency table of the numbers of citations.

The definition of fold enrichment can be written in three different ways but giving identical values, reflecting the different aspects of interpreting the enrichment. First, if we read Table 1 horizontally, the fold enrichment denoted as r can be defined as the ratio of the following two fractions:

$$r = \frac{k/n}{m/N} = \frac{kN}{mn} \tag{1}$$

where k/n is the fraction of B being cited only in A-citing publications, and m/N is the global fraction (or the background probability) of B being cited. When $r > 1$, it means that on average B is cited by A more than B is cited globally. Thus, there is an over-representation of B cited by A compared to the global average. We can conclude that B has a positive influence on A. When $r < 1$, we can conclude that B has a negative influence on A.

Second, if we read Table 1 vertically, r can be defined as the ratio of two fractions in a different form:

$$r = \frac{k/m}{n/N} = \frac{kN}{mn} \tag{2}$$

where k/m is the fraction of citations from A only in B-cited publications, and n/N is the global fraction of citations from A. When $r > 1$, it means that on average A cites B more than A cites globally. Thus, there is an over-representation of A citing B. We conclude that A is positively influenced by B.

Third, we can read Table 1 in both dimensions in a neutral way. We first calculate the following two probability values denoted as p_1 and p_2:

$$\begin{aligned} p_1 &= n/N \\ p_2 &= m/N \end{aligned} \tag{3}$$

where p_1 is the probability of a citation from A and p_2 is the probability of B being cited. If the two events of A citing and B being cited are independent, we can multiply p_1 and p_2 to obtain an expected number of citations from A to B, denoted as k_{exp}:

$$k_{\text{exp}} = p_1 p_2 N = \frac{mn}{N}. \tag{4}$$

Then the fold enrichment r is defined as the ratio of the observed number of citations to the expected number:

$$r = \frac{k}{k_{\text{exp}}} = \frac{kN}{mn}. \tag{5}$$

It is straightforward to see that the three definitions of the fold enrichment r generate the same value. In this study, log2-transformation is additionally applied to the fold enrichment to generate the final metric of enrichment score used in this study, denoted as s:

$$s = \log_2 r. \tag{6}$$

We define that B has a positive influence on A or A is positively influenced by B if $s > 0$, and B has a negative influence on A or A is negatively influenced by B if $s < 0$. If it is necessary to clarify the direction of the influence, we will explicitly add subscripts to the notations such as k_{B-A}, r_{B-A}, s_{B-A}, m_{B-}, and n_{-A} where the first letter before "$-$" corresponds to the cited country and the second letter corresponds to the citing country. If the letter is empty, it means the globe. If A is the same as B, the enrichment score, i.e., s_{A-A}, measures the domestic influence of the country.

The citation network can be constructed as a bipartite graph where citations are as directional edges connecting cited and citing publications. For a universe set of total N citations, i.e., N edges in the bipartite graph, we assume each citation randomly picks one cited publication and one citing publication independently and with equal probability. The random model can also be constructed by randomly permutating the total N edges in the bipartite graph. Then specifically for two countries A and B, the number of citations from A to B is a random variable denoted as X, which follows a hypergeometric distribution $X \sim \text{Hyper}(m, n, N)$ where parameters m, n and N have already been explained in Table 1. The probability of observing a value larger (for measuring the positive influence) or smaller (for measuring the negative influence) than k can be calculated directly from the hypergeometric distribution, or its approximations, the Binomial distribution or chi-square distribution (Rivals et al. 2006). However, in the context of the citation analysis, the numbers of citations in the contingency table are usually huge and it is very likely to obtain a very tiny p-value (e.g., N = 139,760,456 in this study). On the other hand, to reject the null hypothesis that two events of A citing and B being cited are independent is not of interest in this study, while the magnitude of the citation enrichment between the two countries is more relevant. Thus, we only used the enrichment score as the quantitative measurement of the influence between countries while no hypothesis-testing procedure was applied in this study.

4 Results

4.1 International citations are dominant

Before performing the citation enrichment analysis, we first analyzed the fractions of country-wise international citations and references. We restricted to countries with a total number of citations and references both larger than 10,000 in the time period of this study, which resulted in 82 countries and regions[7] covering six continents. It is easy to see in Figure 3A, that the majority of the countries have both high fractions of international citations and references. In 79 out of all 82 (96.3%) countries, more than 74% of their citations

[7]In the remaining part of this paper, they are described only as countries for simplicity.

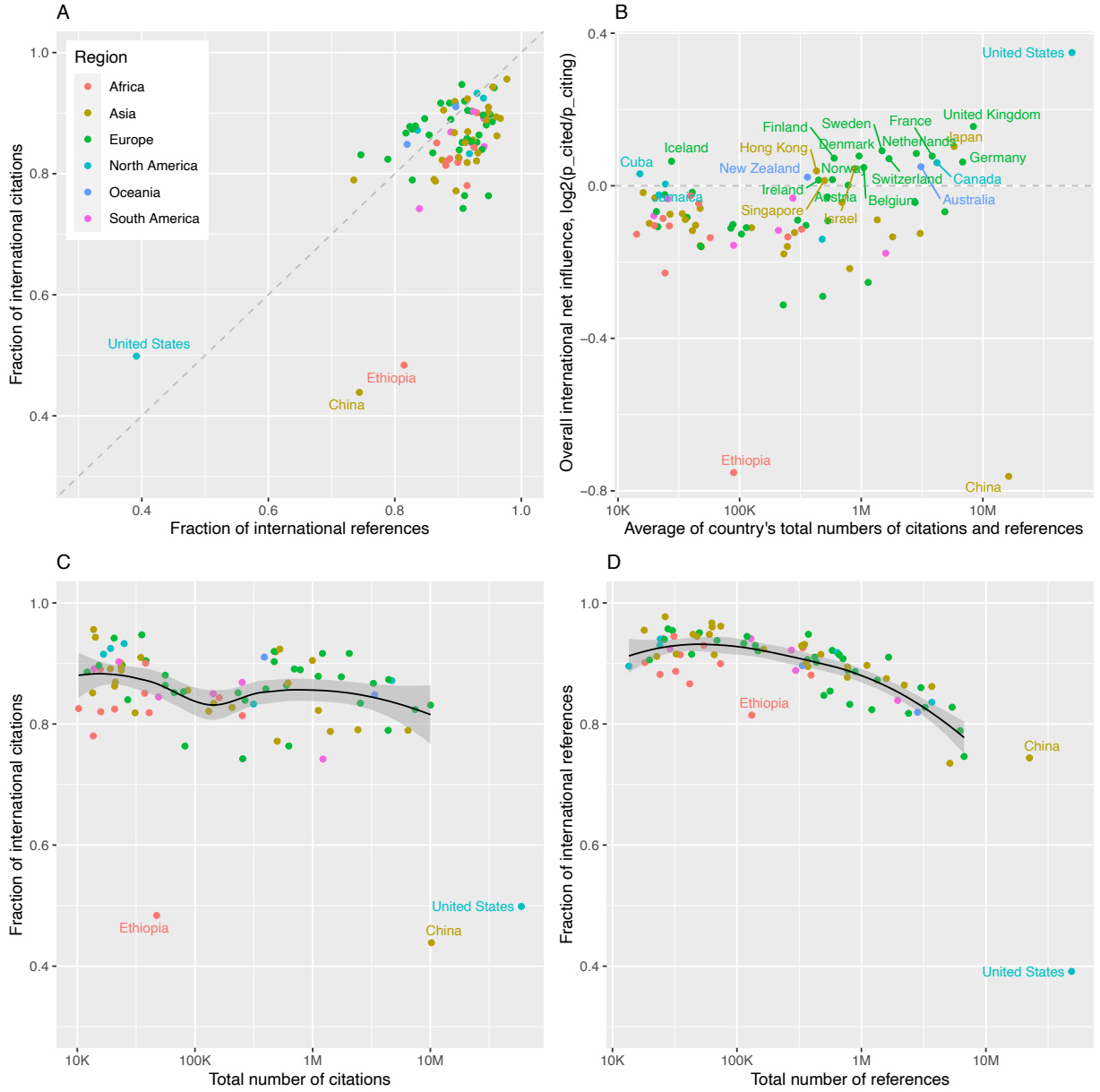


Figure 3: Compare the fractions of international citations and references of 82 countries. A) Scatterplot of the fractions of international references and citations. B) The net international knowledge flows. Countries with values larger than zero are highlighted with their names. C) Relations between the fraction and total number of international references. D) Relations between the fraction and total number of international citations. In Figures C and D, China, Ethiopia and the US were excluded when fitting the loess smoothing lines.

are international. In 81 out of 82 (98.8%) countries, more than 73% of their references are international. There are only three countries with a small fraction of international citations: 43.9% for China, 48.4% for Ethiopia and 49.9% for the US, thus with a relatively high proportion of domestic citations. Only the US has a small fraction (39.1%) of international references.

Denote p_{cited} and p_{citing} as the fractions of international citations and references of a country, we can use $\log_2(p_{\text{cited}}/p_{\text{citing}})$ to measure the "net" flow of how it receives (the in-flow) or contributes (the out-flow) knowledge to the globe. Figure 3B illustrates that countries with positive net out-flow are mostly highly advanced in science (highlighted with labels), whereas the US has the largest net out-flow internationally. Interestingly, China, as the country with the second-largest number of citations in the globe, has the largest net in-flow, implying its global trend of being significantly influenced by international countries.

Excluding China, Ethiopia and the US which behave differently from other countries, we found the fraction of international citations has a weak negative correlation to the absolute number of total citations of the country (Figure 3C). The fraction of international references also has a negative but stronger correlation to the total number of references (Figure 3D). The decrease in international citations and references may be due to that, a high volume of citations is associated with a high volume of publications which in turn is associated with a high volume of scientific units in the country, i.e., more universities and institutes, which eventually promotes more between-units activities within the country such as collaborations and conferences, and increases the fractions of domestic citations.

As has been described in this section, the fraction-based citation analysis can already reveal important patterns of international influence of countries, however, we argue that such analysis is too general and incomplete. Figure 3B shows the US has the largest net out-knowledge flow internationally. But, we should be careful when concluding that the US is the most significant country that contributes scientific influence to the world. In Section 4.3 where we applied citation enrichment analysis on individual country pairs, we demonstrated that the mutual influences between the US and all other countries in the world are actually negative. The net out-flow of the US is due to that the US has a stronger negative preference for citing than being cited by other countries. In other words, the US is cited by other countries less than random but it cites other countries even more less than random. The scenario for China is more complex. With the citation enrichment analysis, we revealed that China's overall net in-flow is mainly caused by the western countries where China is cited less than it cites but the mutual influence between them is negative, while it shows positive influences (i.e., more citations than random) partially on the other parts of the world, To sum up, a single aggregated metric on country-level is not enough for representing the patterns of how the country interacts with the globe especially when the interactions are heterogeneous, which will cause insufficient and inaccurate interpretations. Additionally, with the citation enrichment analysis in Section 4.2, we will give an accurate and quantitative explanation of the negative relations found in Figure 3C-D.

From the next section, we will demonstrate the advantages of citation enrichment analysis for uncovering a detailed and full view of the worldwide scientific influence.

4.2 Domestic citations are highly enriched

We calculated the citation enrichment score from one country to the other with a minimum of 100 citations in the time period of this study. This resulted in 4,583 country pairs covering 82 countries. Note the country pairs are directional. Figure 4A illustrates the global distribution of the citation enrichment scores. In general, the range of enrichment scores decreases when the number of citations increases. This is a typical pattern of the distribution of log2 fold enrichment where the absolute number of observations follows the hypergeometric distribution (Supplementary File 1).

We first studied the enrichment of domestic citations, i.e., how a country cites itself. As shown in Figure 4A, domestic citations are very isolated and show significantly higher enrichment than international citations. Such high enrichment is intuitive to understand because domestic scientific exchanges are expected to be more active than international ones, e.g., via conferences, invited talks, workshops and collaborations within the country, which leads to the highest preference for domestic citations. As described in Section 3.1, if the domestic influence is simply measured by the fraction of domestic citations to the total number of citations or references of that country, as shown in Figure 4B, the ranking of domestic citations drops when the number of domestic citations decreases. In other words, the level of domestic citations will be much underestimated for small countries. Particularly, the US is the most cited country of all countries, and the most referenced country of 51 countries, and China is the most referenced country of 20 countries. This implies that the US or China will always be the most influenced or influencing country if taking the domestic citation fraction as the metric, which is just purely due to their large citation volumes. As an opposite example, the domestic

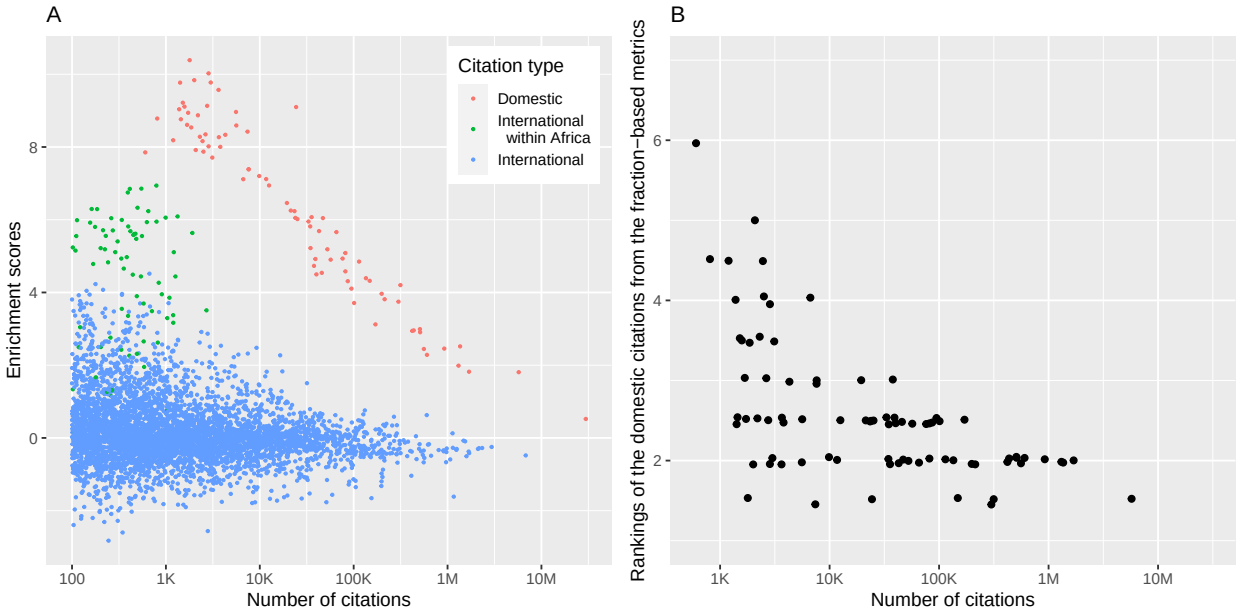


Figure 4: The global citation enrichment. A) The scatterplot of the enrichment scores against the numbers of citations. Each dot corresponds to a citation relation from one country to the other. B) Rankings of the domestic citations from the fraction-based metrics. Values on the y-axis are the average rankings from the cited countries and citing countries. Each dot corresponds to a country.

citation of Macau only has an average ranking of 6 by the fraction metrics[8] due to its very small citation volume, but it has a large enrichment score of 7.85 for its domestic citations, representing a very strong preference for citing and being cited by the publications also produced in Macau. Thus, the enrichment method successfully highlights the cognitive patterns in the domestic citation process, and it reveals the true meaningful patterns of scientific influence.

In Figure 4A, it is easy to see that the enrichment score has a well-linear relation to the number of domestic citations in the log scale (the red dots). For domestic citations, we first found that empirically the number of total citations or references has a linear relation to the number of domestic citations in the log scale (Supplementary File 2), which may be due to the scale-free attribute and the exponential growth of the citation network (Radicchi and Castellano 2015). We first fitted the following two linear models:

$$\begin{aligned} \log_2 n_{-A} &= a_1 \cdot \log_2 k_{A-A} + b_1 \\ \log_2 m_{A-} &= a_2 \cdot \log_2 k_{A-A} + b_2 \end{aligned} \tag{7}$$

where k_{A-A} is the number of domestic citations of country A, n_{-A} is the total number of references of A, and m_{A-} is the total number of citations of A. The estimated coefficients are $a_1 = 0.81$, $b_1 = 6.20$, $a_2 = 0.89$, and $b_2 = 4.45$. The p-values from the F-tests of the two linear regressions are both significant ($p < 2.2 \times 10^{-16}$, Supplementary File 2).

Let's denote T_{A-A}, T_{-A} and T_{A-} as the three citation networks constructed by A-domestic, A-citing and A-cited citations where T_{-A} and T_{A-} can be thought of as two background networks of T_{A-A}. Since the regression coefficients from the linear fits are constants for all involved countries, they can be treated as intrinsic attributes of the citation networks and countries with different sizes correspond to different stages in the citation network expansion. The linear fits in Equation 7 can be transformed to $k_{A-A} \sim n_{-A}^{1/a_1} = n_{-A}^{1.23}$ and $k_{A-A} \sim m_{A-}^{1/a_2} = m_{A-}^{1.12}$. Both power coefficients larger than one implies that T_{A-A} grows faster than its two background networks T_{-A} and T_{A-}, thus increasing fractions of domestic references and citations. Also, we can have $n_{-A} \sim m_{A-}^{a_1/a_2} = m_{A-}^{1.10}$, implying the citing network T_{-A} grows faster than the cited

[8]The average of the ranking from all cited countries and the ranking from all citing countries.

network T_{A-}, thus a faster increasing fraction of domestic references than citations. These results explain the patterns observed in Figure 3C and 3D.

Then the domestic enrichment score can be written as

$$\begin{aligned} s_{A-A} &= \log_2 k_{A-A} - \log_2 n_{-A} + \log_2 N - \log_2 m_{A-} \\ &= (1 - a_1 - a_2) \log_2 k + \log_2 N - b_1 - b_2 \\ &= -0.70 \log_2 k_{A-A} + 16.41 \end{aligned} \tag{8}$$

which is the linear regression fit for domestic citations in Figure 4A.

Let's simplify the problem to $n_{-A} = m_{A-}$, then in the random scenario where two events of "A citing itself" and "A being cited by itself" are independent, with Equation 5 and 6, we can have $a_1 = 0.5$ and $a_2 = 0.5$, and the growth speed of T_{A-A} is quadratic to its background networks, i.e., $1/a_1 = 2$ and $1/a_2 = 2$. When there is a positive domestic enrichment, more citations from the two background networks are added to T_{A-A}, Note the expansion of background networks is approximately linear, which slows down the growth speed of T_{A-A} to $1/a_1 < 2$ and $1/a_2 < 2$, and results in $a_1 + a_2 > 1$, which eventually results in the negative trend between the enrichment scores s_{A-A} and $\log_2 k_{A-A}$ (Equation 8). A complete proof can be found in Supplementary File 2.

Interestingly, we also found there is a group of green dots between the domestic citations and the main dot cloud in Figure 4A. They are citations, although not domestic, but only between African countries. This implies that influences between African countries are much stronger than other within-continent international citation relations.

As domestic citations have distinct enrichment patterns from international ones, we excluded them from the analysis in the next sections and we only explored the international citation relations between countries (green box in Figure 1, but removing the dashed arrows).

4.3 Landscape of the international scientific influence

In Section 4.1, we have demonstrated the scientific knowledge flows in the majority of countries are international, which supports the necessity and importance of studying international influence between countries. To construct the global map of pairwise country-level influences, enrichment scores were formatted into a square matrix and visualized by a comprehensive heatmap in Figure 5. In it, rows correspond to the cited countries and columns correspond to the citing countries, thus influences are transmitted from row-countries to column-countries. Since not every pair of countries had enough citations for calculating the enrichment scores, to reduce the sparsity of the matrix, we additionally required each country to have enrichment scores available to at least other 24 countries. Finally, a 72×72 square matrix with a sparsity of 22.6% was generated for constructing the world influence map[9]. Note the heatmap is not symmetric, although visually very close to. The mutual influence between two countries, i.e., s_{A-B} and s_{B-A}, will be discussed in Section 4.4.

The 72 countries included on the heatmap were partitioned into seven groups both on rows and columns. For the partitioning, we constructed a graph by taking the enrichment matrix as the adjacency matrix and the enrichment scores as the weight of edges. Next, the Louvain graph clustering method (Blondel et al. 2008; Csardi and Nepusz 2006) was applied to partition the graph by the connectivity of nodes, i.e., countries are more densely connected within clusters than between clusters. Since the Louvain method does not allow negative values, we only considered country-pairs with positive enrichment scores. On the heatmap, rows and columns are further reordered by hierarchical clustering within each group. The separation of the seven country groups can also be validated by the dimension reduction analysis on the influence matrix (Supplementary File 3).

Very surprisingly, the partitioning of 72 countries fits perfectly to their geographical classifications and cultural closeness (the "Group" annotation on the heatmap). Countries in the seven groups as well as their geographical classifications are listed in Table 2.

In Figure 5 as well as in Table 2, the first two groups contain countries in northern and western Europe, developed English countries, and Israel. The common attribute of these countries is that they are highly science-advanced and have closer scientific communications compared to other countries. Countries in both groups can be classified into the so-called highly developed "western countries". The US shows a slightly

[9]Note the domestic citations were also removed from this matrix.

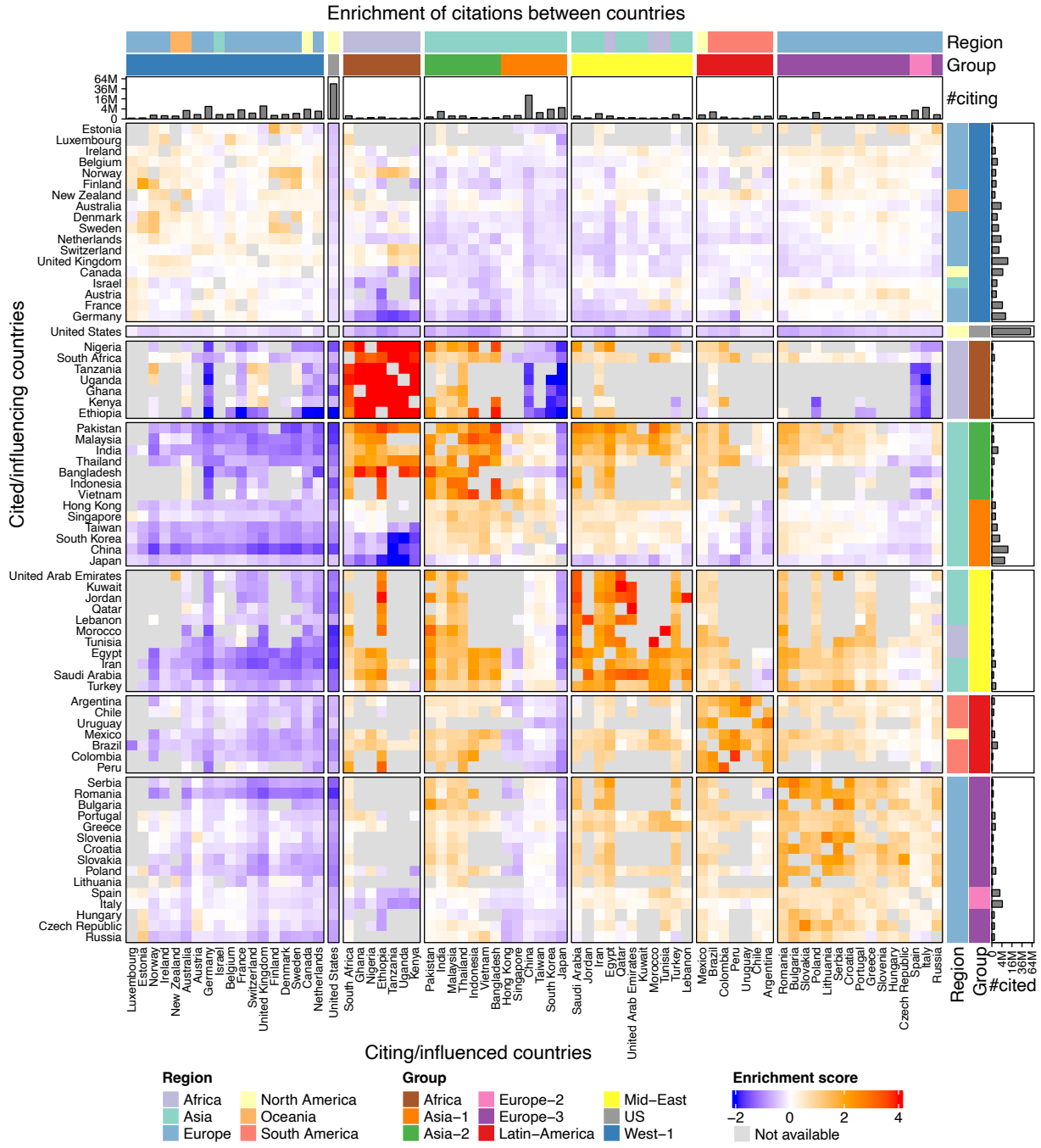


Figure 5: Global map of international scientific influences. Barplots on the two sides of the heatmap are in quadratic scales.

Countries	Geographical regions	Groups
Australia, Austria, Belgium, Canada, Denmark, Estonia, Finland, France, Germany, Ireland, Israel, Luxembourg, Netherlands, New Zealand, Norway, Sweden, Switzerland, United Kingdom	North America, Western Europe, Northern Europe, Oceania, Israel	West-1
United States	North America	US
Ethiopia, Ghana, Kenya, Nigeria, South Africa, Tanzania, Uganda	Africa	Africa
China, Hong Kong, Japan, Singapore, South Korea, Taiwan	East Asia	Asia-1
Bangladesh, India, Indonesia, Malaysia, Pakistan, Thailand, Vietnam	South Asia, Southeast Asia	Asia-2
Egypt, Iran, Jordan, Kuwait, Lebanon, Morocco, Qatar, Saudi Arabia, Tunisia, Turkey, United Arab Emirates	Middle East	Mid-East
Argentina, Brazil, Chile, Colombia, Mexico, Peru, Uruguay	Latin America	Latin-America
Italy, Spain	Southern Europe	Europe-2
Bulgaria, Croatia, Czech Republic, Greece, Hungary, Lithuania, Poland, Portugal, Romania, Russia, Serbia, Slovakia, Slovenia	Eastern Europe, Southern Europe	Europe-3

Table 2: Country groups identified from the world scientific influence map.

overall mutual negative influence on all group-1 countries, thus we only termed group-1 countries as "West-1", but treated the US as a separated group, emphasizing its uniqueness. However, both West-1 and the US show consistent and strong mutual negative influences on all other countries, implying the scientific recognition between the two parts of the world is weaker than random and the globe is split into two isolated worlds. Although West-1 countries in general show mutual negative influence on the remaining countries on the map, there are still exceptions where individual western countries affect other countries positively. For example, Switzerland, Norway and the UK have an intermediate mutual positive influence on the three eastern African countries of Uganda, Kenya and Tanzania ($s = 0.76$)[10], reflecting their international exchange and cooperation of science (Christine and Falko 2022). France has an intermediate positive influence on Tunisia and Morocco ($s = 0.72$), reflecting their cultural closeness and historical relations.

The third group contains seven African countries. They show the strongest mutual positive influences ($s = 5.18$) compared to the within-group influence in all other groups. The high influence within African countries can also be observed in Figure 4A where the green dots are located away from the main dot cloud of international influence. However, the numbers of citations which Africa cites or is cited by other countries are too small that the majority of the enrichment scores are not available for them (light grey grids on the heatmap), reflecting their low level of international scientific influence.

The fourth group contains Asian countries. It can be easily observed from the influence pattern on the heatmap that they can be further split into two subgroups. The first subgroup contains China, Hong Kong, Japan, Singapore, South Korea and Taiwan, which are more science-advanced with larger citation volumes in the Asian regions, but relatively lower levels of mutual positive influence ($s = 0.37$). Note these countries are also culturally similar. Thus, we termed this subgroup as "Asia-1". Even within the Asia-1 group, we can still see that Hong Kong, Japan and Singapore behave slightly differently from China, South Korea and Taiwan where the former three countries are weakly negatively influenced by the Middle East and Latin America and are thought to be closer to western countries than the latter three ones, but the difference is small compared to the global trend, thus Asia-1 was not further split on the map. The second subgroup

[10]The value of s in the parentheses is the average enrichment score from the included countries. It is the same in the remaining part of the paper.

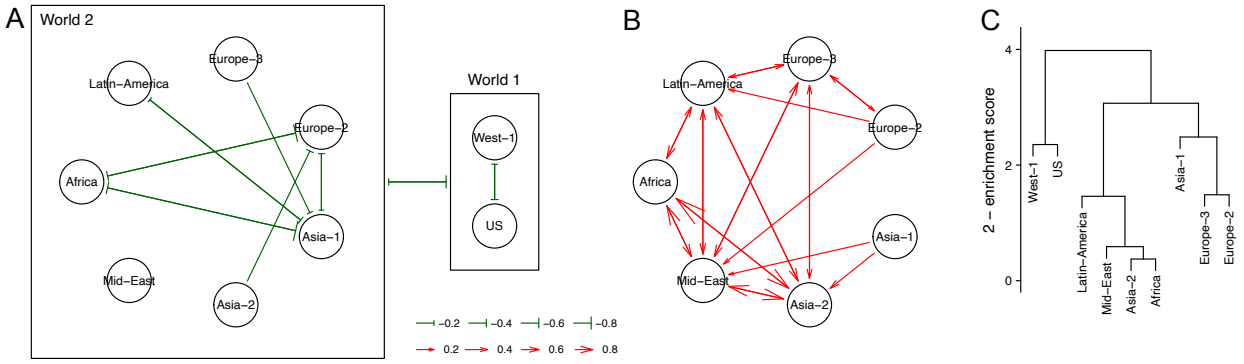


Figure 6: Simplified world scientific influence network. A) The negative influence network. B) The positive influence network. C) Hierarchical relations of country groups based on their pairwise influences. A subtree with a height less than 2 can be interpreted as the average mutual influence of its member country groups being positive.

contains countries in South Asia and Southeast Asia which are less science-advanced in the regions but show higher levels of mutual positive influence ($s = 2.48$). We termed this subgroup as "Asia-2".

The fifth and sixth groups are countries in the Middle East and Latin America. They show consistent mutual positive influence within groups ($s = 2.34$ and 1.80).

The last group contains countries from southern and eastern Europe. They are also manually split into two subgroups based on the patterns on the heatmap. We termed Spain and Italy as "Europe-2" and other group-7 countries as "Europe-3". It might be unexpected that Spain and Italy are thought to be highly developed European countries, but under the context of citation enrichment, they indeed show a negative influence on West-1 ($s = -0.17$) and a positive influence on the Middle East and Europe-3 ($s = 0.33$ and 0.51), making them separated from the western world. This result reveals the imbalance of science within the European Union. The heatmap also reveals Spain has a stronger mutual positive influence on Latin America than Italy ($s = 0.56$ and 0.06), which may relate to the historical relations of Spain and Latin America. Interestingly, we found Spain has a slightly stronger positive influence on Spanish-speaking Latin American countries ($s = 0.59$) than Brazil (Portuguese-speaking, $s = 0.37$), while a higher positive influence was found for Portugal on Brazil ($s = 0.86$). These results reflect the tight associations between the scientific influences and cultural similarity.

4.3.1 Between-group influence: two separated worlds

To better summarize the worldwide influence map, we constructed a simplified group-level influence network. The influence between the two groups was calculated as the average enrichment scores of their member countries. To highlight the strong influence both in the positive and negative directions, we only considered the group pairs with absolute influence larger than 0.15. To reduce the complexity of interpreting the influence network, we separated it into a negative and a positive influence network, as shown in Figures 6A and 6B.

The nine country groups can be put into two categories which we termed "World-1" and "World-2". The World-1 category contains West-1 and the US, which show consistent mutual and strong negative influences on the countries in World-2 ($s = -0.74$ for World-2 on World-1 and $s = -0.30$ for World-1 on World-2, Figure 6A). In World-1, the US shows an intermediate level of mutual negative influence on West-1 countries ($s = -0.35$ for West-1 on the US and $s = -0.28$ for the US on West-1), but it is mainly due to the high domestic citations in the US.

The remaining countries compose World-2 and they are quite separated from World-1. In general, countries in World-2 are more mutually positively influenced. Especially, Africa, Asia-2 and Mid-East show the strongest mutual positive influences ($s = 2.20$, Figure 6B). There are still regional negative influences sourced from Europe-2 and Asia-1. Interestingly, Europe-2 and Asia-1 show certain levels of exclusive enrichment patterns in World-2: Europe-2 has a mutual positive influence on Europe-3 ($s = 0.40$) and contributes a positive influence on Latin-America ($s = 0.31$) while Asia-1 is negative on both ($s = -0.22$ and -0.26); Asia-1 shows

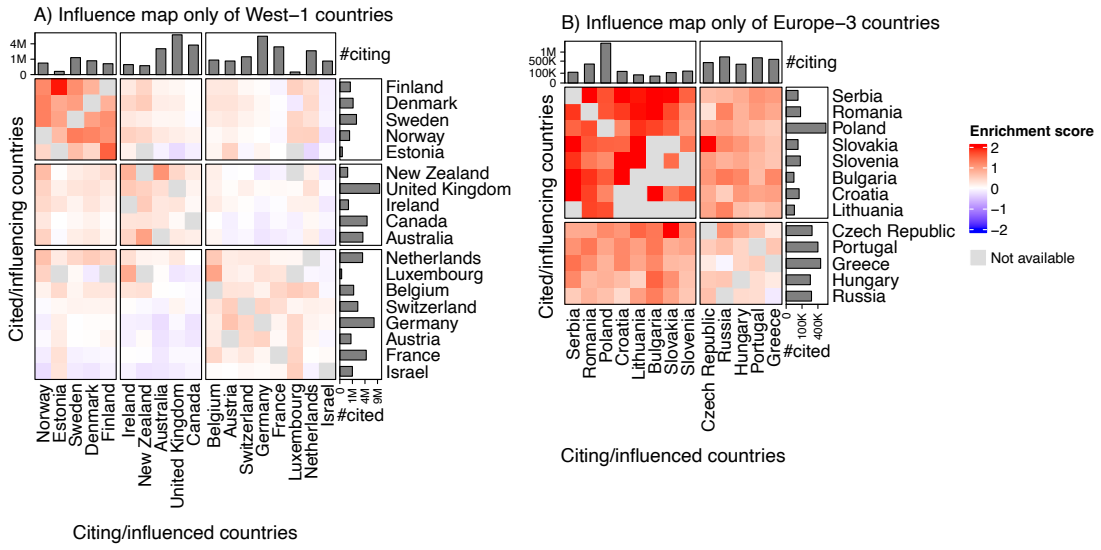


Figure 7: International influcence of A) 18 West-1 countries and B) 13 Europe-3 countries. Barplots on the two sides of the heatmap are in quadratic scales.

positive influence on Asia-2 ($s = 0.56$) while Europe-2 receives a negative influence from it ($s = -0.39$). The mutual influence between Europe-2 and Asia-1 themselves is $s = -0.24$. However, Europe-2 and Asia-1 show positive influences both on Mid-East ($s = 0.33$ and 0.23) and mutual negative influence both on Africa ($s = -0.86$ and -1.09).

It is interesting to see, Europe-2 and Asia-1, which are more science-advanced in World-2, contribute negative influences in the regions. Together with the separation of World-1 from World-2, we could infer that the development of science may slow down the spread of the scientific influence on the regions that are weakly related to, which separates them apart on the world influence map.

Figure 6C illustrates the hierarchical relations of the nine country groups. In the hierarchical clustering, the distance matrix D was generated from the group-level average enrichment scores where a distance value d_{ij} between group i and j is calculated as:

$$d_{ij} = 2 - \text{sign}(s_{ij} + s_{ji}) \cdot \max(|s_{ij}|, |s_{ji}|) \tag{9}$$

where the value is subtracted from 2 to make the distance value positive. The dendrogram confirms the separation of World-1 and World-2, the clustering of Europe-2 and Europe-3, and the clustering of Latin-America, Mid-East, Asia-2 and Africa.

4.3.2 Within-group influence: small worlds

Within individual country-groups, we also observed "small world" patterns, where countries inside can be further split into subgroups. As already shown in Figure 5, by directly observing the influence patterns on the heatmap, Asian countries are split into Asia-1 and Asia-2, and southern and eastern European countries are split into Europe-2 and Europe-3.

West-1 shows a certain level of heterogeneity of mutual enrichment. Similar to the partitioning analysis of the global map, we applied the Louvain method only on the 18 West-1 countries, generating three subgroups both on rows and columns (Figure 7A). We found that countries also exhibit geographical and cultural similarities within the subgroups. The first subgroup contains the five northern European countries which show stronger mutual positive influence ($s = 1.05$) than other West-1 countries. The second subgroup contains five English countries ($s = 0.40$). The third subgroup mainly contains mainland European countries ($s = 0.29$, excluding Israel). The northern Europe has a consistent mutual positive influence on English countries ($s = 0.19$). Interestingly, in mainland Europe, Netherlands, Luxembourg and Belgium show relatively stronger positive mutual influence on northern Europe ($s = 0.26$) and English countries ($s = 0.12$), as their geographical locations as bridges connecting the regions. The other four mainland European countries, i.e., France,

Germany, Switzerland and Austria, show almost-neutral mutual influence on English countries ($s = -0.07$) and northern Europe ($s = 0.06$). Last, it can be observed that Estonia receives positive influence from mainland Europe ($s = 0.23$), reflecting the geographical closeness between them. Such a small-world pattern of countries is mainly observable within West-1. The country-subgroups can not be well distinguished from how they interact with World-2 countries (Supplementary File 4).

Although countries in Europe-3 are overall mutually positively influenced, a certain level of heterogeneity can still be observed. In Figure 7B, since almost all mutual influences are positive, making the graph clustering approach less effective, we partitioned countries simply by hierarchical clustering. We found that there is a core set of 8 countries (Serbia, Romania, Poland, Slovakia, Slovenia, Bulgaria, Croatia and Lithuania) showing very strong mutual influence ($s = 1.80$) and they have consistent smaller citation volumes (excluding Poland) than other countries in Europe-3.

4.4 Balance of mutual citation enrichment

The country-to-country influence is directional. A country contributes influences to other countries and also receives influences from them. Specifically for two countries A and B, the mutual influences are s_{A-B} and s_{B-A}. If A influences B in the same amount as being influenced by B, measured by $s_{A-B} = s_{B-A}$, we could say that the mutual influence between A and B is balanced. If $s_{A-B} > s_{B-A}$, we say A has a net out-influence on B or else A receives a net in-influence from B. As we have shown that countries in the same country group share very similar enrichment patterns, also to reduce the complexity of the interpretation, in this section, we studied the influence balance in each country group only by selecting a representative country. Additionally, by observing the balance patterns, Asia-1 is additionally split into Asia-1-1 (China, South Korea, and Taiwan) and Asia-1-2 (Hong Kong, Japan, and Singapore). Africa was removed from the analysis because the enrichment scores are missing between it and most of the other countries. We selected the following nine countries: The US, Netherlands (West-1), Spain (Europe-2), Poland (Europe-3), Brazil (Latin-America), Saudi Arabia (Mid-East), Thailand (Asia-3), China (Asia-1-1) and Hong Kong (Asia-1-2). Figure 8 illustrates the mutual influences of the selected countries to all individual countries. The complete analysis for all individual countries can be found in Supplementary File 5.

In Figure 8, most data points are located close to the line of $y = x$, which implies that in general countries are approximately mutual influence-balanced. This pattern can also be observed from Figure 5 where the lower triangle shows similar patterns as the upper triangle on the square heatmap. We found, that for each country in Figure 8, mutual influence patterns are approximately linear over countries, and the linear relations are distinct between World-1 and World-2. In the citation balance analysis, we took $y = x$ (completely balanced) as the baseline, thus, we first fitted the ordinary linear model of $y-x \sim y+x$, which is identical to minimizing the sum of the square distances of data points to $y = x$ in the fit, then the fitted equation was transformed to the form of $y = \alpha x + \beta$.

To better explain the slope coefficient α, let's take country A as a fixed country and $B \in W$ where W is the set of countries in World-1 or World-2. The linear model between s_{A-B} and s_{B-A} is

$$s_{A-B} = \alpha \cdot s_{B-A} + \beta. \tag{10}$$

Simply speaking, α corresponds to the increasing rate of s_{A-B} against s_{B-A}. When $\alpha > 1$, s_{A-B} increases faster than s_{B-A}, i.e., A influences B faster than B influences A. Let's assume there are two countries B_1 and B_2 with $s_{A-B_2} > s_{A-B_1}$, i.e., A has a stronger enrichment on B_2 than B_1. With Equations 5, 6 and 10:

$$\begin{aligned} \frac{k_{A-B_2}N}{m_{A-}n_{-B_2}} &= \left(\frac{k_{B_2-A}N}{m_{B_2-}n_{-A}}\right)^{\alpha} 2^{\beta} \\ \frac{k_{A-B_1}N}{m_{A-}n_{-B_1}} &= \left(\frac{k_{B_1-A}N}{m_{B_1-}n_{-A}}\right)^{\alpha} 2^{\beta}. \end{aligned} \tag{11}$$

Dividing Equations 11 on both sides yields:

$$\begin{aligned} \frac{k_{A-B_2}/n_{-B_2}}{k_{A-B_1}/n_{-B_1}} &= \left(\frac{k_{B_2-A}/m_{B_2-}}{k_{B_1-A}/m_{B_1-}}\right)^{\alpha} \\ \frac{p_{A-B_2}}{p_{A-B_1}} &= \left(\frac{p_{B_2-A}}{p_{B_1-A}}\right)^{\alpha}. \end{aligned} \tag{12}$$

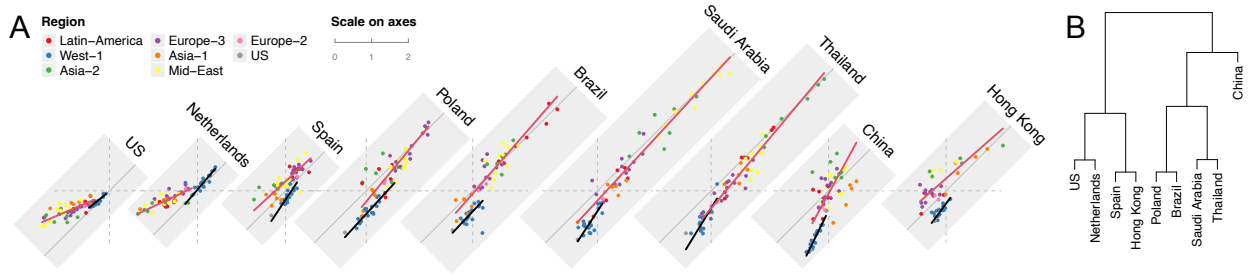


Figure 8: Influence balance of the nine selected countries. A) Each grey rectangle corresponds to the plotting panel of a country. s_{B-A} corresponds to values on the x-axis and s_{A-B} corresponds to values on the y-axis where A corresponds to one of the nine selected countries and B corresponds to individual countries in each plotting panel. All nine panels are aligned by their x-axes and have parallel y-axes. The scales on the x-axes and y-axes are the same for all panels. The linear fit is applied to World-1 and World-2 separately. B) Hierarchical clustering on the citation balance patterns in the two worlds. Africa was removed from the analysis.

Denote four citation networks as T_{A-B}, T_{B-A}, T_{-B} and T_{B-} that are constructed by citation relations in the four categories where $B \in \{B_1, B_2\}$, then k_{A-B}, k_{B-A}, n_{-B} and m_{B-} are the sizes of their citation networks. T_{-B} can be thought of as the background network of T_{A-B}, and T_{B-} as the background network of T_{B-A}. Then, e.g., p_{A-B_1} is the fraction of T_{A-B_1} in its background network T_{B_1} measured by their sizes. The condition $s_{A-B_2} > s_{A-B_1}$ implies $p_{A-B_2} > p_{A-B_1}$, thus, we can define the relative growth rate of T_{A-B} from B_1 to B_2 denoted as $\rho_{T_{A-B}}$ as:

$$\rho_{T_{A-B}} = p_{A-B_2}/p_{A-B_1} > 1. \tag{13}$$

According to Figure 8, the slope coefficients for all linear regressions are larger than zero, then with Equation 10 and $\alpha > 0$, we also have $s_{B_2-A} > s_{B_1-A}$ then $p_{B_2-A} > p_{B_1-A}$. This means when A gives more influence on B, at the same time it also receives more influence from B. Then the relative growth rate of T_{B-A} from B_1 to B_2 denoted as $\rho_{T_{B-A}}$ can be similarly defined as:

$$\rho_{T_{B-A}} = p_{B_2-A}/p_{B_1-A} > 1. \tag{14}$$

Equation 12 becomes:

$$\rho_{T_{A-B}} = \rho^{\alpha}_{T_{B-A}}. \tag{15}$$

Note we already have $\rho_{T_{A-B}} > 1$ and $\rho_{T_{B-A}} > 1$, then we can have the following relations:

$$\begin{aligned} \rho_{T_{A-B}} &> \rho_{T_{B-A}} \quad \text{if } \alpha > 1, \\ \rho_{T_{A-B}} &< \rho_{T_{B-A}} \quad \text{if } \alpha < 1. \end{aligned} \tag{16}$$

Equation 16 implies, when $\alpha > 1$ and when the citation enrichment of A increases from B_1 to B_2, the relative growth of T_{A-B} (or the growth of the fraction of T_{A-B} from its background network T_{-B}) is faster than its reverse citation network T_{B-A}, and vice versa. Thus, when there is no ambiguity, we say A has a faster mutual influence expansion on B when $\alpha > 1$, and A has a slower mutual influence expansion on B when $\alpha < 1$.

The nine selected countries have well-linear mutual influence relations to the countries in World-1 or World-2, then, we can treat e.g., World-1 as a system where its individual countries correspond to different stages but in the same speed of influence expansion. Under this context, α can be generalized from country-to-country to a system-level metric of the mutual influence of A and a world W. A larger value of α represents a faster mutual influence expansion of A on W.

When $\alpha \approx 1$, the intercept coefficient β can be calculated as:

$$\beta \approx s_{A-B} - s_{B-A} = \log_2 \left(\frac{k_{A-B}}{k_{B-A}} \Big/ \frac{k_{A-B,\exp}}{k_{B-A,\exp}} \right) \tag{17}$$

where k_{A-B}/k_{B-A} is the ratio of the numbers of mutual citations between A and B, and $k_{A-B,\exp}/k_{B-A,\exp}$ is the mutual citation ratio assuming A and B citing each other randomly. Thus, β can be approximately considered as the log fold enrichment of the mutual citation between two countries. Note in this inference, A is a fixed country and β is a constant for any $B \in W$, thus the estimated coefficient β can be considered as an intrinsic attribute of A on W, which corresponds to the baseline of mutual influence to all other countries in a world. When $\beta > 0$, we say A and W are systematically mutually influenced, when $\beta < 0$, we say A and W are systematically mutually exclusive, and when $\beta = 0$, we say A and W are systematically balanced. A detailed explanation of β can be found in Supplementary File 6.

The overall influence balance of A denoted as η_A is simply defined as:

$$\eta_A = \frac{1}{n_W} \sum_{B \in W} s_{A-B} - s_{B-A} \tag{18}$$

where n_W is the number of countries in W. A positive value of η_A represents the net out-influence of A on the respective world.

The coefficients of the linear fits as well as the overall influence balance for the nine countries on World-1 and World-2 are listed in Table 3. In Figure 3B, we have demonstrated the overall trend of the US is exporting scientific knowledge to the world. Here with the citation balance analysis, more accurate internal structures can be revealed. First, in Figure 8 it is straightforward to see, the net out-influence of the US is mainly from World-2 ($\eta = 0.56$) where corresponding countries are all above $y = x$, while it is almost balanced on West-1 ($\eta = 0.07$). As the slope coefficient is far less than 1 ($\alpha = 0.43$) on World-2, it implies the US receives the scientific influence from World-2 faster than it contributes the influence. More specifically, the relative growth of the citation network of the US citing World-2 scaled by its background network from World-2 is faster than that of the reverse citation network of the US cited by World-2. Second, all the enrichment scores between the US and World-2 are negative, implying the overall net out-influence of the US is due to that, the citing activities between the US and World-2 are both weaker than random, but the US cites World-2 even less than it is cited. Third, notice that the intercept coefficient between the US and World-2 is close to zero ($\beta = -0.06$), then, we can conclude that, although the US cites World-2 weaker than it is cited, as the US influences World-2 more, i.e., from high negative to weak negative, the process of the US influencing World-2 expands faster than the reverse process, which eventually makes the mutual influence between them approaching balanced.

China is another extreme country in Figure 3B which receives the largest in-knowledge flow from the world. Here as revealed in Figure 8, the citation influence balance pattern for China is more complex, where it has a stable net in-influence flow from World-1 but a heterogeneous and partial net out-influence on World-2. On World-1, China has the largest slope coefficient ($\alpha = 2.04$) in the nine countries, implying the influence from China to World-1 increases much faster than World-1 on China, although the citation process between China and World-1 is still weaker than random. The intercept coefficient $\beta = -0.54$ implies a systematic mutual exclusion of the influence between World-1 and China. The balance between China and World-2 is very heterogeneous and there is no reliable linear relation of the mutual influence if taking World-2 as a whole (a negative R^2 of -0.51). However, if inspecting individual country groups, we can still observe China has a net in-influence from Asia-1, but out-influence on Asia-2. There are also linear mutual influences between China and Asia-1, Mid-East and Europe-3 separately. The three slope coefficients ($\alpha = 1.99$, 1.93 and 2.96, Supplementary File 7) imply China has a faster mutual influence expansion on countries in these three groups. The intercept coefficient of China on Asia-1 ($\beta = -1.18$, Supplementary File 7) implies a systematic mutual influence exclusion between them.

Netherlands which represents West-1 behaves very similarly to the US since they all belong to the same category of World-1. It has a positive intercept coefficient ($\beta = 0.14$) on World-2 compared to the US, implying the mutual activity between World-2 and Netherlands is systematically stronger than random.

Interestingly, under the context of influence balance analysis, We found, that although Spain and Hong Kong are categorized into World-2, they show reverse mutual influence patterns to other World-2 countries. When they interact World-2, we found they have net out-influence ($\eta = 0.25$ and 0.35) while other selected

	Country group	World-1				World-2				In-group
		α	β	η	R^2	α	β	η	R^2	η
United States	US	0.82	-0.01	0.07	0.70	0.43	-0.06	0.56	0.48	-
Netherlands	West-1	1.17	0.04	0.09	0.88	0.47	0.14	0.55	0.59	0.10
Spain	Europe-2	1.50	-0.24	-0.21	0.75	0.92	0.26	0.25	0.60	0.07
Hong Kong	Asia-1-2	1.34	-0.35	-0.39	0.62	0.88	0.34	0.35	0.63	-0.04
Poland	Europe-3	1.09	-0.63	-0.60	0.79	1.21	-0.29	-0.14	0.71	-0.04
Brazil	Latin-America	1.06	-0.56	-0.61	0.72	1.15	-0.08	0.02	0.84	0.03
Saudi Arabia	Mid-East	1.53	-0.56	-0.67	0.44	1.08	-0.27	-0.17	0.94	-0.10
Thailand	Asia-2	1.63	-0.47	-0.57	0.19	1.20	-0.47	-0.28	0.93	-0.07
China	Asia-1-1	2.04	-0.54	-0.93	0.86	1.83	-0.07	-0.02	-0.51	-0.50

Table 3: Coefficients of the linear fits as well as the overall influence balance in the citation balance analysis. α: slope coefficient, β: intercept coefficient, η: overall influence balance, R^2: R-squared statistic from the linear regression. The R^2 of the linear regression of China on World-2 is negative because the fitting is not directly from ordinary linear regression, and it also implies a highly unreliable fit. The last column: the overall influence balance of the country to other countries within the same group. The splitting of countries is based on the clustering in Figure 8B.

World-2 countries have close-to-zero or overall in-influence to World-2. Spain and Hong Kong also preferably mutually influence World-2 ($\beta = 0.26$ and 0.34) while other four selected World-2 countries show certain levels of exclusivity. These results imply that Spain and Hong Kong are active knowledge contributors in World-2. Spain and Hong Kong also have positive intercept coefficients ($\beta = 0.26$ and 0.34) on World-2, reflecting the positive mutual influence between them and World-2 on the baseline level. When interacting with World-1, we found it is a common pattern that all selected World-2 countries have net in-influence from World-1 (the values of η are all negative) and mutually exclusive influence on the system-level (the values of β are all negative). However, Spain and Hong Kong have weaker imbalance and exclusivity from World-1 than other World-2 countries ($\beta = -0.24$ and -0.35 verse -0.55, $\eta = -0.21$ and -0.39 verse -0.68)[11].

The remaining four countries (Poland, Brazil, Saudi Arabia and Thailand) have similar balance patterns. Particularly, Poland and Brazil have slop coefficients close to 1 both on World-1 ($\alpha = 1.09$ and 1.06) and World-2 ($\alpha = 1.21$ and 1.15), implying their stable expansions of scientific influence over the world.

We also investigated the overall influence balance of a country within its own group. For example, the average mutual influence balance of Netherlands on all other 17 West-1 countries is 0.10. As shown in the last column of Table 3, mutual scientific influence is almost balanced inside all country groups, except Asia-1-1 which is mainly because China has a huge imbalance of citing and being cited by international publications (Figure 3A-B). However, South Korea and Taiwan which are also in Asia-1-1, are still very mutually balanced ($\eta = 0.02$).

Figure 8B illustrates the relations of the nine countries by hierarchical clustering on the regression coefficients from linear fits on World-1 and World-2. Compared to the dendrogram in Figure 6C which is based on the static world influence patterns, in the citation balance analysis which additionally considers the dynamical relation of a country and a world, we found Poland, Brazil, Saudi Arabia and Thailand, which represent Europe-3, Latin-America, Mid-East and Asia-2, are still tightly clustered. However, Hong Kong and Spain, which represent partial Asia-1 and Europe-2, are now clustered together and closer to World-1. This implies, that although statically, Hong Kong and Spain fall within World-2 due to their relatively lower fold citation enrichment to World-1, but their dynamics of interacting the globe are closer to World-1. The separation of Spain and Hong Kong from World-1 is more from a systematic imbalance.

[11] -0.55 and -0.68 are the average of the β and η coefficients of Poland, Brazil, Saudi Arabia, Thailand and China on World-1.

5 Discussion

Using citation data, the scientific influence among countries can be defined on the macro-scale, such as the total number of international citations of a country, which behaves like background influence where a country with a large portion plays a major role in the globe. The scientific influence can also be measured on the micro-scale, e.g., the average number of international citations on the single publication level. In this study, instead of considering the absolute or relative amount of international citations, we have proposed a new metric "citation enrichment" which is constructed on another layer where citations are randomly made between publications. With it, we defined the scientific influence from a new aspect as how preferably a country cites publications from another country, and it correctly reflects the cognitive behavior of researchers in the citation process. The citation enrichment can also be thought of as a relative metric defined on the micro-scale, but it corrects the system-level difference caused by citation volumes of countries. Additionally, citation enrichment is a country-to-country metric, which extends the country-wise citation analysis into two dimensions, and it greatly improves the deciphering of internal structures of the complex citation relations over countries.

We have demonstrated the significance of the citation enrichment metric by an extensive analysis of over 14 million publications in the life science and biomedical fields. We successfully revealed that domestic scientific influence is the strongest in a country compared to all its international influences, and such enrichment is not hidden by the citation volumes of small countries. This prioritization of domestic influence aligns perfectly with the pattern of science dissemination where domestic science activity is believed to be the strongest. The domestic influence result agrees with Bakare and Lewison (2017) who defined an "Over-Citation Rate" in the same form as the citation enrichment, but applied at the publication level.

For the first time, we have constructed a comprehensive world map of scientific influence among 72 countries across six continents utilizing the citation enrichment metric. This influence map has a highly structured layout where countries are clustered into groups based on their mutual positive and negative influences. The perfect associations of these country groups with the geographical and cultural relations among their member countries reflect the selective nature of scientific exchange between countries. We revealed the general structure of global science influence where developed countries are isolated from the rest of the world. Even within World-2, the two groups of Europe-2 and Asia-1-2, which are relatively science-advanced in the regions, exhibit mutual exclusivity from each other while their positive influences are only observed in their neighbouring regions. This interesting negative association between a country's level of scientific development and its influence on other regions in the world suggests a slowing scientific dissemination as scientific advancement increases, and scientific knowledge is more concentrated within the country or its local neighbourhood. This phenomenon might be an interesting topic for further exploration and validation.

Citation is a directional relation. To comprehensively understand a country's role in the global science dissemination, it is important to study the bi-directional influence between the country and the globe. Thus, we performed the influence balance analysis and studied the patterns of how a country cites and how it is cited. An interesting finding is, that most countries, regardless of their regions of origin, have well-linear relations of mutual influences in the two worlds separately. Thus, we could say each world ensembles a system where its member countries have similar mechanisms of mutually influencing others. With it, we can study the dynamics of a country interacting with a world, i.e., how the mutual influence changes across countries in a world, the baseline and the overall trend. If the influence map describes relations between countries from static influence patterns, the influence balance analysis studies the worldwide scientific influence from another aspect of how a country dynamically influence other countries.

From the aspect of methodology, the enrichment score measures the association between two events represented in the row and column of the 2×2 contigency table and restricted in a certain universe set. By choosing different events and universes, the enrichment analysis can be generalized as a powerful tool for bibliometrics analysis. Let's denote the row and column events as E_a and E_b, and the universe set as U. In this study, E_a is "country A cites", E_b is "country B is cited", U is the set of total citations in the globe. As examples, we can choose other events for E_b to study international influences on various levels:

- Journal-to-country: E_b is "a journal is cited".
- Institution-to-country: E_b is "an institution is cited".
- Author-to-country: E_b is "an author is cited".
- Publication-to-country: E_b is "a publication is cited".

Then specifically, we can answer the question "does a journal/institution/author/publication have general international influence or is its influence only restricted to one or a few countries?"

Under the same framework, the citation enrichment metric can also be applied in longitudinal studies, e.g., in the following two directions:

- Define E_a as "country A cites", E_b as "country B is cited only for its publications published on year y", and U as the set of total citations on and after year y. By applying the citation enrichment analysis to a list of continuous years, we can study the longitudinal trend of country-to-country influences, and answer the question "are the two worlds becoming more separated?"
- Define E_a as "country A cites on year y_i", E_b as "country B is cited only for its publications published on year y_j", and U as the set of total citations of publications that are published on year y_j and cited on y_i. Of course we have $y_i \geq y_j$. Then we can study the short-term influence if y_i and y_j are close and the long-term influence if y_i and y_j are far apart.

The 2×2 contingency table in Table 1 contains numbers of citations in different categories. This means the basic element is a citation in the analysis. The contingency table can also be constructed from the number of publications. In this way, E_a is "a publication is produced by authors from country A", and E_b is "a publication is produced by authors from country B". Then k in Table 1 measures the number of publications produced by authors from both A and B, i.e., via collaborations from the two countries. This means, the citation enrichment analysis can be extended to the publication level to study the levels of international collaborations, and eventually to construct a collaboration world map.

The universe set U is also an important component in the framework of citation enrichment analysis from two aspects. First, all the categories in the 2×2 contingency table should be restricted within the universe set, thus, the selection of the universe affects the values in the contingency table. Second, the enrichment score is a relative metric against a random process which is performed within the universe. The universe defines *what are the citations enriched against* and it should have a clear meaning for the interpretation. In this study, we defined U as the set of the total citations among the selected countries. A large value of s measures a stronge enrichment of the citations between two countries against a random citation process where publications randomly cite each other regardless whether they are from the same countries or not. With this universe set, we observed that the US has an overall negative influence even on West-1 countries because of its high volume of domestic citations. To get rid of such effect, we can define U as the set of total citations but only international ones. Note for two countries A and B, all the values except k in Table 1 need to be recalculated to remove the counts from domestic citations. As shown in Supplementary File 9, the US now has a mutual positive influence on all West-1 countries. However, note that the interpretation has been changed. Now the positive influence represents more citations than random international citations (i.e., a publication picks a random international publication to cite). It only measures international citation behaviors and conclusions should only be made under that context. Such interpretation is isolated from domestic citations. Last, U can be set for a subset of regional countries, e.g., all citations in European publications and to only study the scientific influence in Europe.

There is one point of the citation enrichment analysis that needs to be careful when interpreting the results. As shown in Figure 4A, although the expected value of the enrichment score is zero, there is still a dependency between its variation and the number of citations. This results in, if there is a meaningful enrichment, either positive or negative, countries with larger citation volumes are more difficult to obtain larger absolute enrichment scores. Thus, it is suggested to compare citation enrichment between countries with similar citation volumes. The citation enrichment metric proposed in this study simply divides the observed number of citations from the expected one without considering the variance if treating the number of citations as a random variable X. There is another widely used method in statistics that measures the relative distance of an observed value of X to its null distribution, the z-score, defined as $z = (k - k_{\text{exp}})/\sigma$ where σ is the standard deviation of X from the hypergeometric distribution. In Supplementary File 6, we demonstrated that using z-scores generates very similar country classifications as in Section 4.3, but the variance of the z-scores now has a positive dependency on the number of citations. Nevertheless, the form of fold enrichment is simpler and more convenient for further statistical analysis.

Similar forms of citation enrichment have been mentioned in literature. On the publication level, Balland and Boschma (2022) and Fontana, Montobbio, and Racca (2019) applied a metric named "Relative Scientific Advantage" to measure whether publications from a country are enriched in a specific research topic, i.e., to evaluate country-topic associations. A positive enrichment is called a specialization of the country on the research topic. Finardi and Buratti (2016) applied a metric named "Probability Affinity Index" to study the levels of collaboration between two countries. Nevertheless, the framework of the enrichment analysis we proposed in this study is more general and extensible, and it can be applied to a wide range of research topics.

Finally, this study is based on the literature data from the PubMed database. PubMed is publically free with no restriction on obtaining its resources. However, it is limited to bibliometrics studies because it mainly contains publications in the biomedical fields. Nevertheless, with the scripts we have provided, the same studies can be easily applied to datasets from other commercial bibliographic databases such as Scopus[12] or World of Knowledge[13] to extend to the full spectrum of academic research fields.

6 Conclusion

We have proposed a new metric, the citation enrichment, under a well-defined statistical framework. With the citation enrichment metric, we constructed a global scientific influence map covering major countries in the world, which uncovers the details of the exchange of biomedical scientific knowledge across borders. The citation balance analysis reveals the dynamics of mutual influence. In this study, we demonstrated that the citation enrichment metric has a clear meaning for interpretation. The enrichment method is also extensible to a general tool. We believe it will be a powerful tool for citation analysis to generate more interesting results in related fields.

CRediT authorship contribution statement

Zuguang Gu: Conceptualization; Formal analysis; Investigation; Methodology; Visualization; Writing - original draft; Writing - review & editing.

Declaration of interest

Declarations of interest: none.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Abramo, Giovanni, Ciriaco Andrea D'Angelo, and Massimiliano Carloni. 2019. "The Balance of Knowledge Flows." *Journal of Informetrics* 13 (1): 1–9. https://doi.org/https://doi.org/10.1016/j.joi.2018.11.001.

Aria, Massimo, and Corrado Cuccurullo. 2017. "Bibliometrix: An r-Tool for Comprehensive Science Mapping Analysis." *Journal of Informetrics* 11 (4): 959–75. https://doi.org/https://doi.org/10.1016/j.joi.2017.08.007.

Bakare, Victoria, and Grant Lewison. 2017. "Country over-Citation Ratios." *Scientometrics* 113 (2): 1199–1207. https://doi.org/10.1007/s11192-017-2490-z.

Balland, Pierre-Alexandre, and Ron Boschma. 2022. "Do Scientific Capabilities in Specific Domains Matter for Technological Diversification in European Regions?" *Research Policy* 51 (10): 104594. https://doi.org/https://doi.org/10.1016/j.respol.2022.104594.

Blondel, Vincent D, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. 2008. "Fast Unfolding of Communities in Large Networks." *Journal of Statistical Mechanics: Theory and Experiment* 2008 (10): P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008.

Bogart, Chris, Christian Kästner, James Herbsleb, and Ferdian Thung. 2021. "When and How to Make Breaking Changes: Policies and Practices in 18 Open Source Software Ecosystems." *ACM Trans. Softw. Eng. Methodol.* 30 (4). https://doi.org/10.1145/3447245.

Carammia, Marcello. 2022. "A Bibliometric Analysis of the Internationalisation of Political Science in Europe." *European Political Science* 21 (4): 564–95. https://doi.org/10.1057/s41304-022-00367-9.

Chinchilla-Rodríguez, Cassidy R. AND Larivière, Zaida AND Sugimoto. 2019. "Follow the Leader: On the Relationship Between Leadership and Scholarly Impact in International Collaborations." *PLOS ONE* 14 (6): 1–18. https://doi.org/10.1371/journal.pone.0218309.

[12] https://www.scopus.com/

[13] https://www.webofknowledge.com/

Christine, Estreguil, and Buschke Falko. 2022. "The Evolving Role of the European Commission in Research on Africa." Scientific analysis or review KJ-NA-31-195-EN-N (online). Luxembourg (Luxembourg): Publications Office of the European Union. https://doi.org/10.2760/38335%20(online).

Csardi, Gabor, and Tamas Nepusz. 2006. "The Igraph Software Package for Complex Network Research." *InterJournal* Complex Systems: 1695. https://igraph.org.

Donthu, Naveen, Satish Kumar, Debmalya Mukherjee, Nitesh Pandey, and Weng Marc Lim. 2021. "How to Conduct a Bibliometric Analysis: An Overview and Guidelines." *Journal of Business Research* 133: 285–96. https://doi.org/https://doi.org/10.1016/j.jbusres.2021.04.070.

Finardi, Ugo, and Andrea Buratti. 2016. "Scientific Collaboration Framework of BRICS Countries: An Analysis of International Coauthorship." *Scientometrics* 109 (1): 433–46. https://doi.org/10.1007/s11192-016-1927-0.

Fontana, Magda, Fabio Montobbio, and Paolo Racca. 2019. "TOPICS AND GEOGRAPHICAL DIFFUSION OF KNOWLEDGE IN TOP ECONOMIC JOURNALS." *Economic Inquiry* 57 (4): 1771–97. https://doi.org/https://doi.org/10.1111/ecin.12815.

Hassan, Saeed-Ul, and Peter Haddawy. 2013. "Measuring International Knowledge Flows and Scholarly Impact of Scientific Research." *Scientometrics* 94 (1): 163–79. https://doi.org/10.1007/s11192-012-0786-6.

Lancho Barrantes, Bárbara S., Vicente P. Guerrero Bote, Zaida Chinchilla Rodríguez, and Félix de Moya Anegón. 2012. "Citation Flows in the Zones of Influence of Scientific Collaborations." *Journal of the American Society for Information Science and Technology* 63 (3): 481–89. https://doi.org/https://doi.org/10.1002/asi.21682.

Radicchi, Filippo, and Claudio Castellano. 2015. "Understanding the Scientific Enterprise: Citation Analysis, Data and Modeling." In *Social Phenomena: From Data Analysis to Models*, edited by Bruno Gonçalves and Nicola Perra, 135–51. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-14011-7_8.

Rivals, Isabelle, Léon Personnaz, Lieng Taing, and Marie-Claude Potier. 2006. "Enrichment or depletion of a GO category within a class of genes: which test?" *Bioinformatics* 23 (4): 401–7. https://doi.org/10.1093/bioinformatics/btl633.

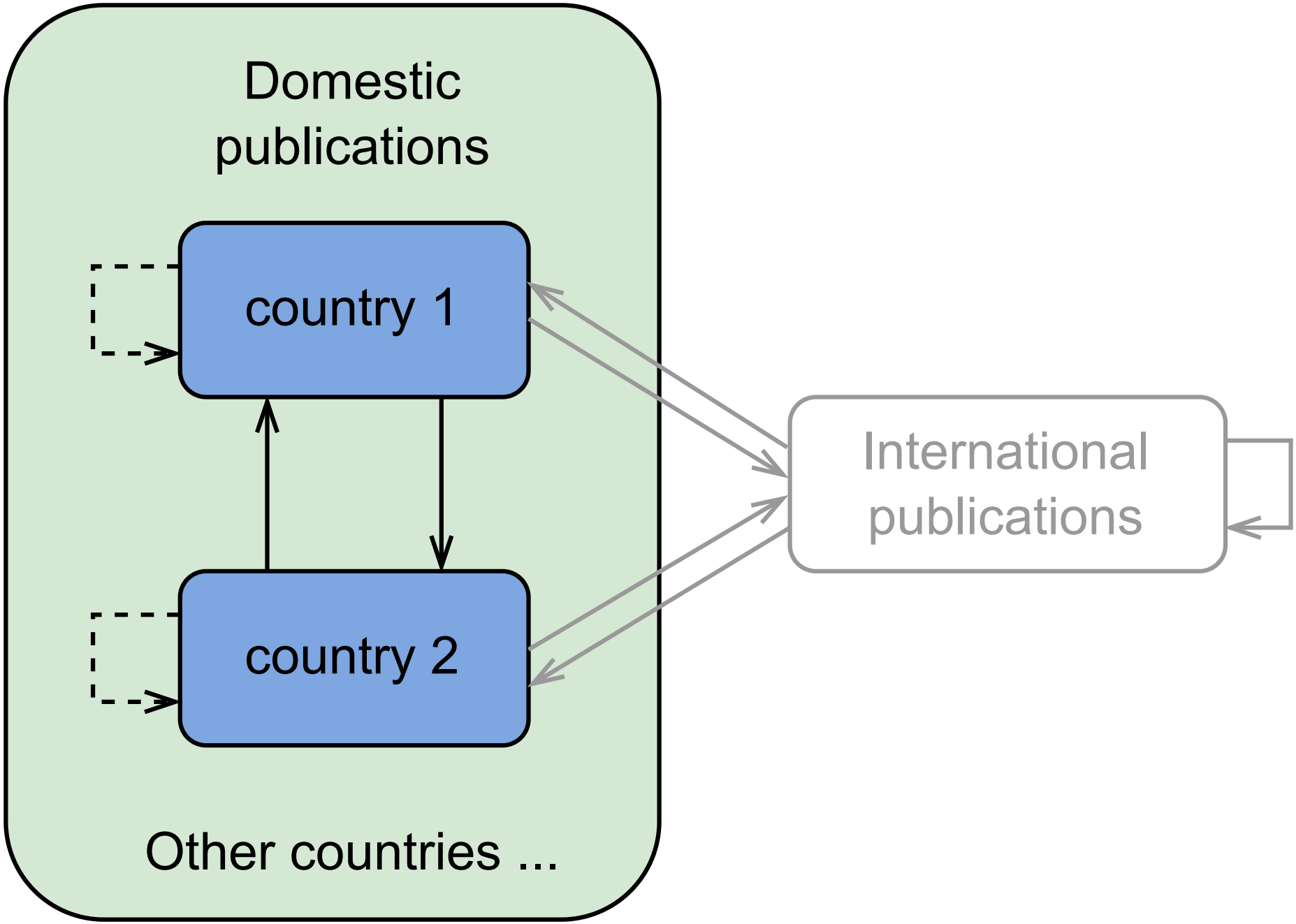
Tahamtan, Iman, and Lutz Bornmann. 2018. "Core Elements in the Process of Citing Publications: Conceptual Overview of the Literature." *Journal of Informetrics* 12 (1): 203–16. https://doi.org/https://doi.org/10.1016/j.joi.2018.01.002.

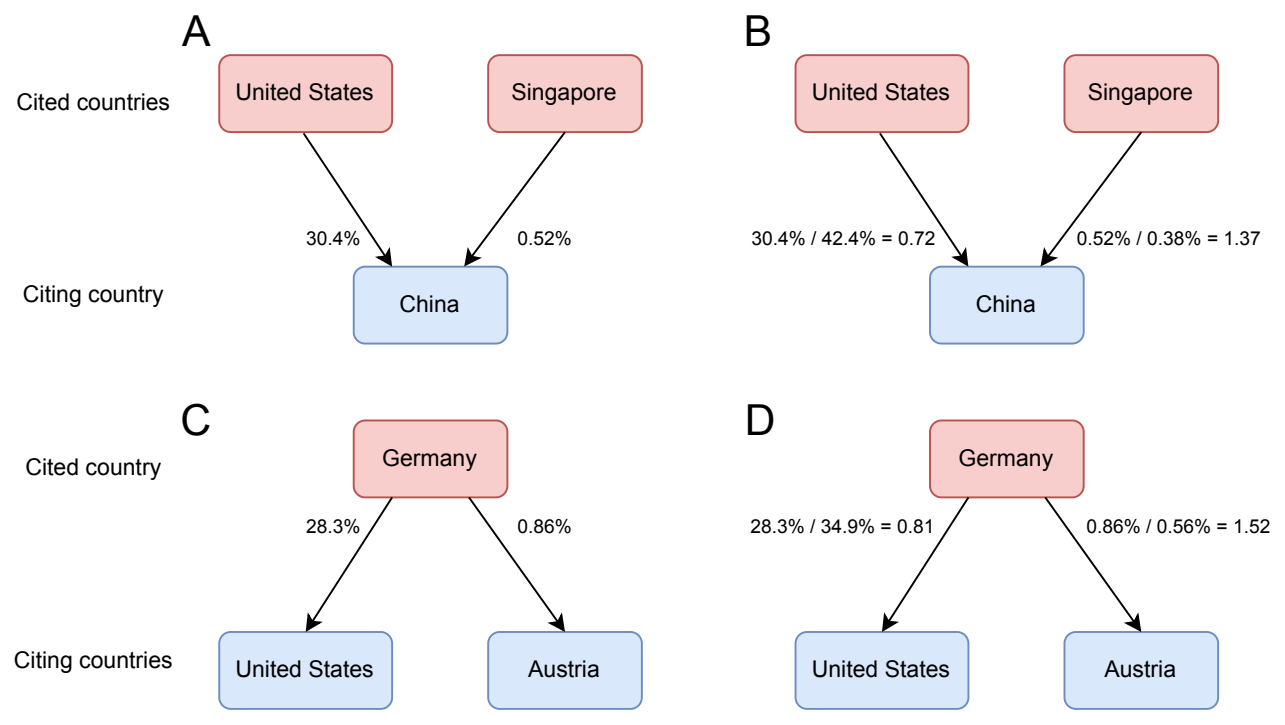
Urlings, Miriam J. E., Bram Duyx, Gerard M. H. Swaen, Lex M. Bouter, and Maurice P. Zeegers. 2021. "Citation Bias and Other Determinants of Citation in Biomedical Research: Findings from Six Citation Networks." *Journal of Clinical Epidemiology* 132: 71–78. https://doi.org/https://doi.org/10.1016/j.jclinepi.2020.11.019.

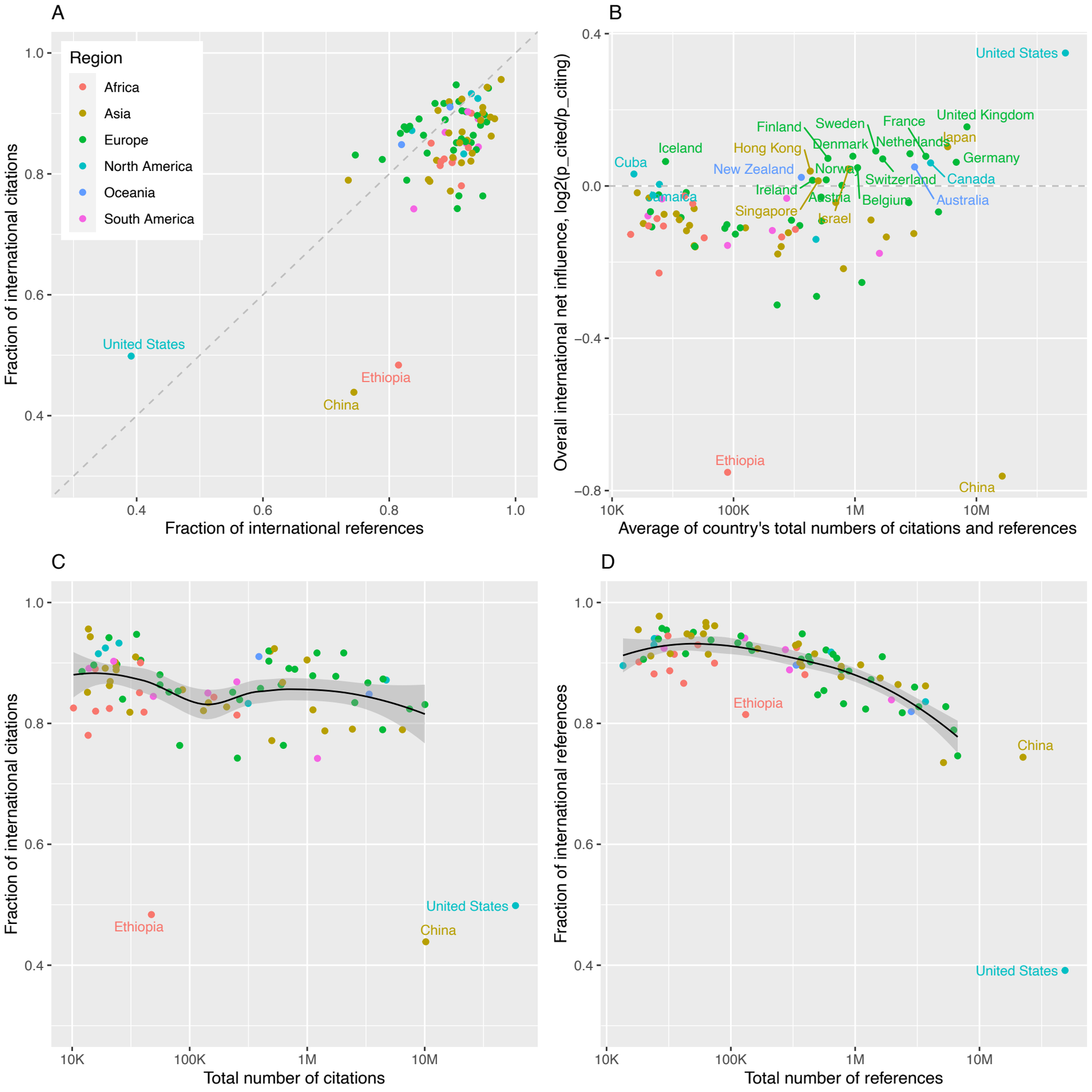
Vieira, Elizabeth S. 2023. "The Influence of Research Collaboration on Citation Impact: The Countries in the European Innovation Scoreboard." *Scientometrics* 128 (6): 3555–79. https://doi.org/10.1007/s11192-023-04715-4.

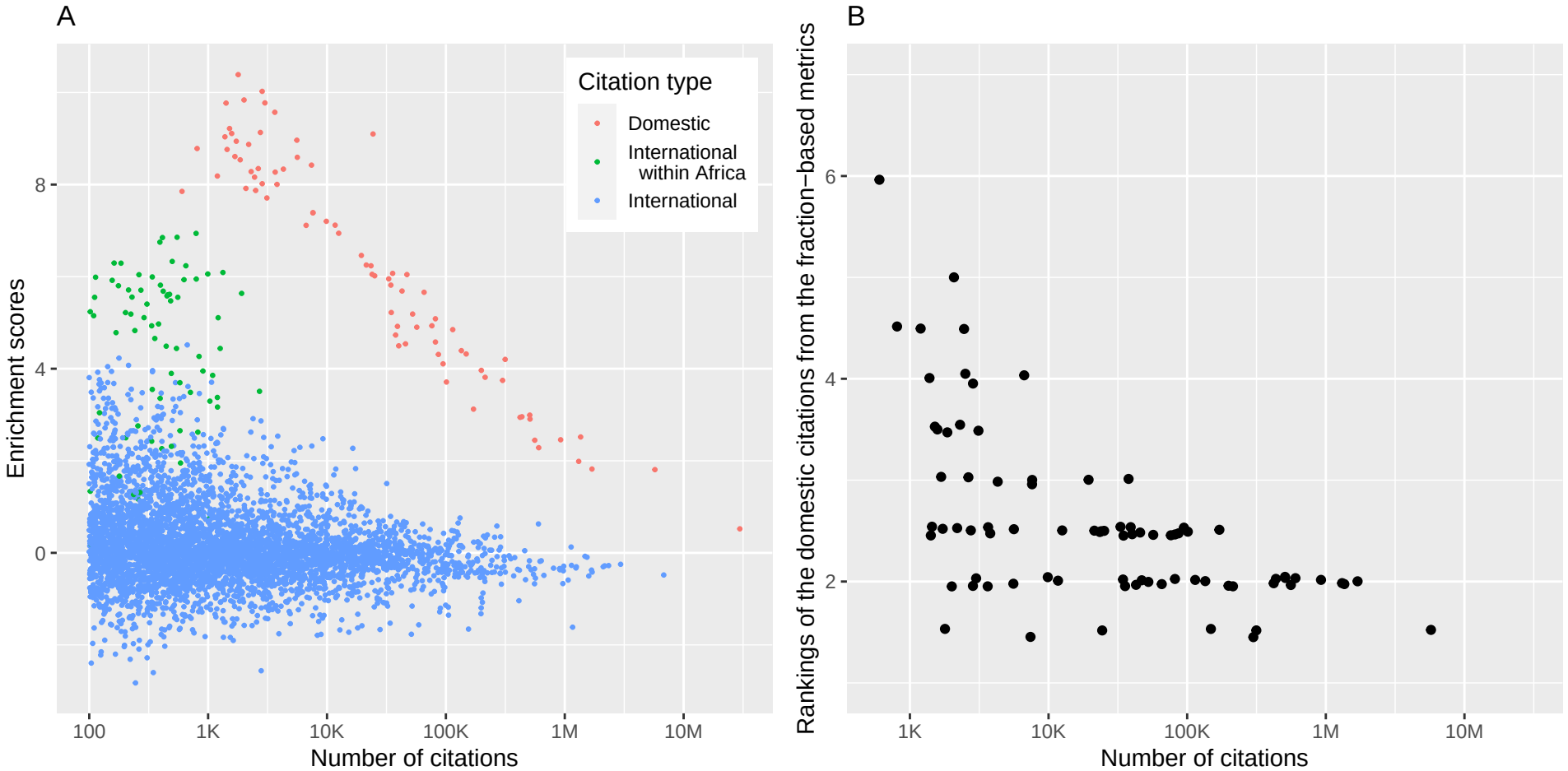
Wren, Jonathan D., and Constantin Georgescu. 2022. "Detecting Anomalous Referencing Patterns in PubMed Papers Suggestive of Author-Centric Reference List Manipulation." *Scientometrics* 127 (10): 5753–71. https://doi.org/10.1007/s11192-022-04503-6.

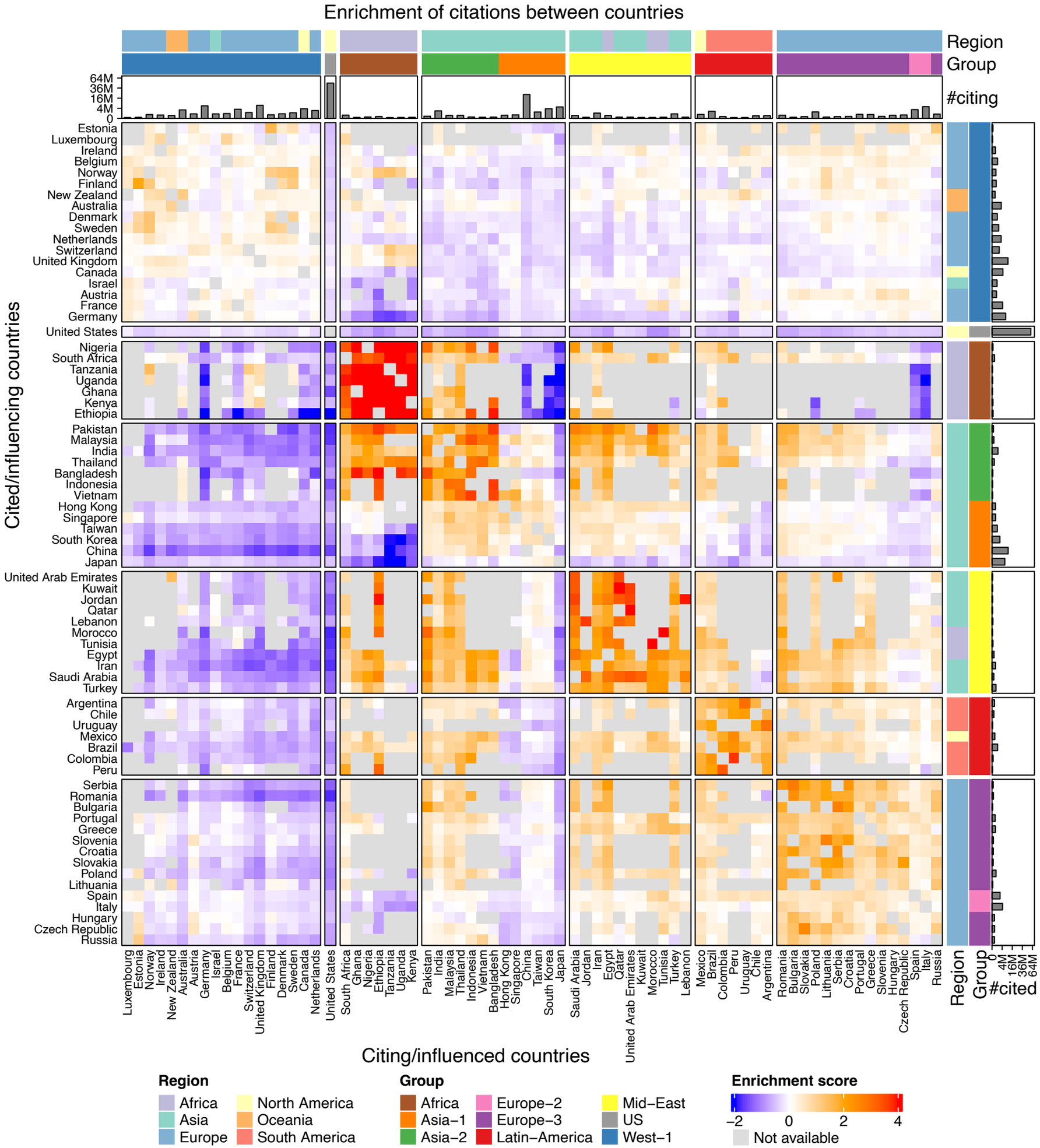
Zhang, Fang, and Shengli Wu. 2021. "Measuring Academic Entities' Impact by Content-Based Citation Analysis in a Heterogeneous Academic Network." *Scientometrics* 126 (8): 7197–7222. https://doi.org/10.1007/s11192-021-04063-1.

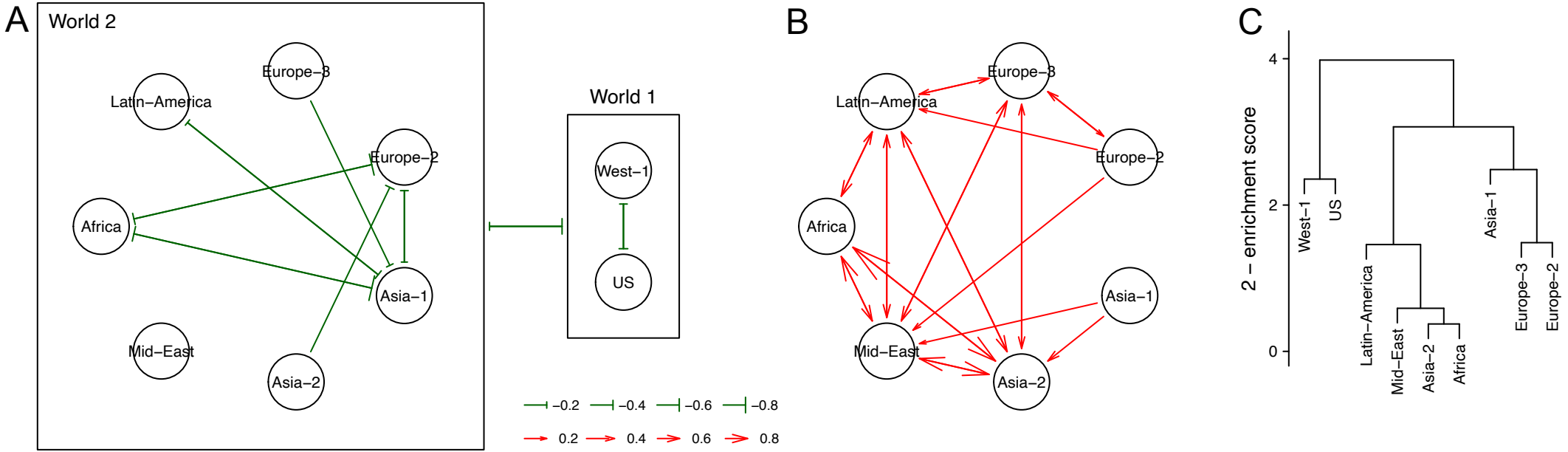

Domestic publications
country 1
country 2
Other countries ...
International publications

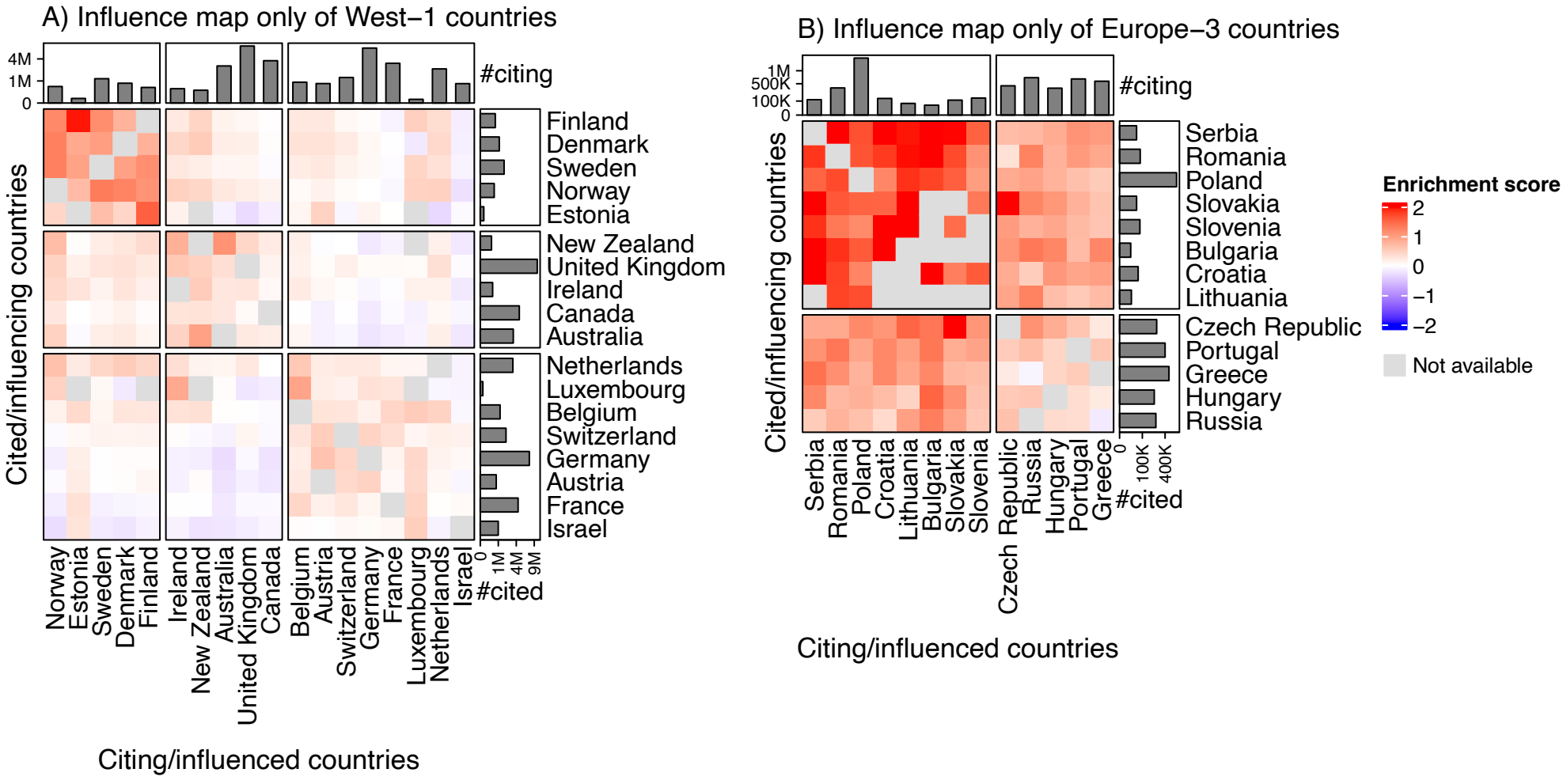

A
Cited countries
United States
Singapore
30.4%
0.52%
Citing country
China
B
United States
Singapore
30.4% / 42.4% = 0.72
0.52% / 0.38% = 1.37
China
C
Cited country
Germany
28.3%
0.86%
Citing countries
United States
Austria
D
Germany
28.3% / 34.9% = 0.81
0.86% / 0.56% = 1.52
United States
Austria

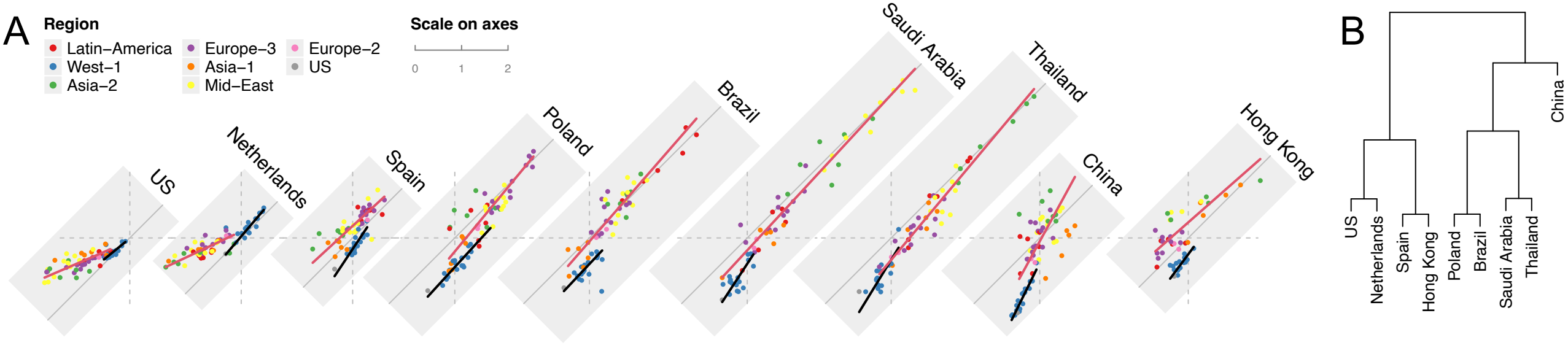

A
Region
Africa
Asia
Europe
North America
Oceania
South America
Fraction of international citations
Fraction of international references
United States
Ethiopia
China
B
Overall international net influence, log2(p_cited/p_citing)
Average of country's total numbers of citations and references
United States
United Kingdom
Finland
Sweden
France
Japan
Iceland
Hong Kong
Denmark
Netherlands
Germany
Cuba
New Zealand
Norway
Switzerland
Canada
Ireland
Jamaica
Austria
Belgium
Australia
Singapore
Israel
Ethiopia
China
C
Fraction of international citations
Total number of citations
United States
Ethiopia
China
D
Fraction of international references
Total number of references
Ethiopia
China
United States


A
Citation type
Domestic
International within Africa
International
Enrichment scores
8
4
0
100
1K
10K
100K
1M
10M
Number of citations
B
Rankings of the domestic citations from the fraction−based metrics
6
4
2
1K
10K
100K
1M
10M
Number of citations


Enrichment of citations between countries
Region
Group
#citing
64M
36M
16M
4M
0
Cited/influencing countries
Estonia
Luxembourg
Ireland
Belgium
Norway
Finland
New Zealand
Australia
Denmark
Sweden
Netherlands
Switzerland
United Kingdom
Canada
Israel
Austria
France
Germany
United States
Nigeria
South Africa
Tanzania
Uganda
Ghana
Kenya
Ethiopia
Pakistan
Malaysia
India
Thailand
Bangladesh
Indonesia
Vietnam
Hong Kong
Singapore
Taiwan
South Korea
China
Japan
United Arab Emirates
Kuwait
Jordan
Qatar
Lebanon
Morocco
Tunisia
Egypt
Iran
Saudi Arabia
Turkey
Argentina
Chile
Uruguay
Mexico
Brazil
Colombia
Peru
Serbia
Romania
Bulgaria
Portugal
Greece
Slovenia
Croatia
Slovakia
Poland
Lithuania
Spain
Italy
Hungary
Czech Republic
Russia
Citing/influenced countries
#cited
Region
Group
Africa
Asia
Europe
North America
Oceania
South America
Africa
Asia−1
Asia−2
Europe−2
Europe−3
Latin−America
Mid−East
US
West−1
Enrichment score
−2
0
2
4
Not available


A
World 2
Europe−3
Latin−America
Europe−2
Africa
Asia−1
Mid−East
Asia−2
World 1
West−1
US
−0.2
−0.4
−0.6
−0.8
0.2
0.4
0.6
0.8
B
Europe−3
Latin−America
Europe−2
Africa
Asia−1
Mid−East
Asia−2
C
2 − enrichment score
4
2
0
West−1
US
Latin−America
Mid−East
Asia−2
Africa
Asia−1
Europe−3
Europe−2


A) Influence map only of West−1 countries
#citing
4M
1M
0
Finland
Denmark
Sweden
Norway
Estonia
New Zealand
United Kingdom
Ireland
Canada
Australia
Netherlands
Luxembourg
Belgium
Switzerland
Germany
Austria
France
Israel
0
1M
4M
9M
#cited
Cited/influencing countries
Norway
Estonia
Sweden
Denmark
Finland
Ireland
New Zealand
Australia
United Kingdom
Canada
Belgium
Austria
Switzerland
Germany
France
Luxembourg
Netherlands
Israel
Citing/influenced countries
B) Influence map only of Europe−3 countries
#citing
1M
500K
100K
0
Serbia
Romania
Poland
Slovakia
Slovenia
Bulgaria
Croatia
Lithuania
Czech Republic
Portugal
Greece
Hungary
Russia
0
100K
400K
#cited
Cited/influencing countries
Serbia
Romania
Poland
Croatia
Lithuania
Bulgaria
Slovakia
Slovenia
Czech Republic
Russia
Hungary
Portugal
Greece
Citing/influenced countries
Enrichment score
2
1
0
−1
−2
Not available


A
Region
Latin−America
West−1
Asia−2
Europe−3
Asia−1
Mid−East
Europe−2
US
Scale on axes
0
1
2
US
Netherlands
Spain
Poland
Brazil
Saudi Arabia
Thailand
China
Hong Kong
B
US
Netherlands
Spain
Hong Kong
Poland
Brazil
Saudi Arabia
Thailand
China