Global disparities in plant science: a legacy of colonialism, patriarchy, and exclusion

Authors: Rose A. Marks[1,2,3,*], Erik J. Amézquita[4,*], Sarah Percival[5,*], Alejandra Rougon-Cardoso[6,7], Claudia Chibici-Revneanu[8], Shandry M. Tebele[3], Jill M. Farrant[3], Daniel H. Chitwood[1,4], Robert VanBuren[1,2]
**contributed equally*

1. *Department of Horticulture, Michigan State University, East Lansing, MI 48824, USA*
2. *Plant Resilience Institute, Michigan State University, East Lansing, MI 48824, USA*
3. *Department of Molecular and Cell Biology, University of Cape Town, Rondebosch 7701, South Africa*
4. *Department of Computational Mathematics, Science, and Engineering, Michigan State University, East Lansing, MI 48824, USA*
5. *Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI 48824, USA*
6. *Laboratory of Agrigenomic Sciences, Universidad Nacional Autónoma de México, ENES-León, León 37689, México*
7. *Plantec National Laboratory, ENES-León, León 37689, México*
8. *Intercultural Development and Management, Universidad Nacional Autónoma de México, ENES-León, León 37689, México*

Keywords: *colonialism, patriarchy, gender equality, eurocentrism, taxonomic gaps, diversity, inclusion*

Running head: *Global disparities in plant science*

Correspondence: *Rose A. Marks, Department of Horticulture, Michigan State University, East Lansing, MI 48824, USA; Email: marksros@msu.edu; Phone: (603) 852-3190; ORCID iD: https://orcid.org/0000-0001-7102-5959*

ABSTRACT

The field of plant science has grown dramatically in the past two decades, but global disparities and systemic inequalities persist. Here, we analyzed ~300,000 papers published over the past two decades to quantify disparities across nations, genders, and taxonomy in the plant science literature. Our analyses reveal striking geographical biases, where affluent nations dominate the publishing landscape and vast areas of the globe have virtually no footprint in the literature. Authors in Northern America are cited nearly twice as many times as authors based in Sub-Saharan Africa and Latin America, despite publishing in journals with similar impact factors. Gender imbalances show remarkably little improvement over time, and some of the most affluent nations have extremely male biased publication records, despite supposed improvements in gender equality. In addition we find that most studies focus on economically important crop and model species and a wealth of biodiversity is under-represented in the literature. Taken together, our analyses reveal a problematic system of publication, with persistent bias that poorly captures the global wealth of scientific knowledge and biological diversity. We conclude by highlighting immediately addressable disparities and suggestions for long-term solutions to improve equity in the plant sciences.

INTRODUCTION

Plant science research is accelerating at an unprecedented pace. New technologies and expanding infrastructure have opened the door for cutting-edge research to be conducted at monumental scales. Despite this noteworthy growth, access to resources is not evenly distributed across the globe and recent studies have revealed striking participation gaps and longstanding disparities tied to colonialism, economic inequality, and systemic biases (Amarante et al. 2021; Maas et al. 2021; Marks et al. 2021; Trisos et al. 2021; Ebenezer et al. 2022; Topaz et al. 2022; Wapman et al. 2022). Plant science, in particular, suffers from more acute historical exclusion and ongoing underrepresentation of marginalized identities compared to other biological disciplines (Madzima and MacIntosh, 2021). In Northern America, associations between plant science and agriculture with colonialism, slavery, and the exploitation of migrant workers (Carter and Alexander, 2020) have contributed to a disproportional lack of diversity in plant science compared to related fields. Global economic disparities, established under imperial colonialism and perpetuated through modern eurocentric frameworks, further exacerbate underrepresentation of diverse perspectives in plant science (Baber 2016; Li 2021; Marks et al. 2021). Researchers working in low-income countries and under-resourced institutions face multiple barriers to participating in plant science research, including limited funding opportunities, reduced access to cutting-edge technologies and infrastructure, and exclusion from collaboration networks (Ebenezer et al., 2022; Ibe, 2022). In the field of plant genomics for instance, few projects have been led by researchers in the Global South, despite the striking biodiversity and extensive local knowledge within these regions (Marks et al., 2021). These dynamics are reinforced by a eurocentric framework that centers English language standards, Latin binomial naming conventions, and reductionist thinking. Coupled with historical and ongoing expropriation of plant germplasm from the Global South, this has resulted in a system that unjustly benefits certain individuals and excludes others. A first step of addressing these inequalities is to quantify patterns of participation in plant science.

Both race and gender compound with global economic disparities to generate emergent barriers for people of color and individuals with marginalized gender identities (Crenshaw, 1989; Maas et al., 2021). For example, women of color are uniquely oppressed across multiple axes in ways that amount to more than the sum of their racial and gender identities (Crenshaw, 1989). Although our analyses do not address race directly, we explore global patterns with links to imperial colonialism that cannot be understood without an acknowledgement of race and the persistent oppression faced by people of color, especially Black and Indigenous communities. Our analyses address patriarchy, sexism, and gender dynamics more directly. Patriarchy can be described as a way of living that privileges all men over women and some men over other men, and the politics of patriarchy can be understood as “the politics of domination – a politics that rationalizes inequality” (Gilligan et al. 2018). Systems of patriarchy vary in their manifestation and severity globally, but are pervasive and have infiltrated all levels of society including scientific research (Haghighat-Sordellini 2010; Golash-Boza et al. 2019; Uchendu and Edeagu 2021; Kocabıçak 2022). While self-identified women are not excluded from the field of biology as a whole, they are often excluded from prestigious tenured and editorial positions as well as collaboration networks (West et al. 2013; Frances et al. 2020; Madzima and MacIntosh 2021; Lerman et al. 2022). Studies suggest that gender biases exist in hiring, publication, and funding decisions (Larivière et al. 2013; Fox et al. 2016; Bonham and Stefan 2017; Holman et al. 2018; Witteman et al. 2019; Frances et al. 2020; Wapman et al. 2022). These inequities impact academic currency on job and funding markets and further exacerbate gender imbalances in academia. Quantifying the extent and patterns of gender bias in plant science is an important step in moving towards a more equitable discipline.

Despite noteworthy efforts made towards cataloging all life, research attention has not been equally distributed across study systems and many species remain underexplored. In plant genomics, for example, there are substantial taxonomic gaps and multiple clades lack a reference genome assembly. (Vallée et al., 2016; Marks et al., 2021; Szövényi et al., 2021). These studies suggest that research attention has been disproportionately directed towards a few select species with agricultural and economic relevance to modern society. Focusing on these elite crop and model species has enabled noteworthy scientific breakthroughs and agricultural innovations, but it has come at the cost of exploring the rich biodiversity of wild plants and regionally important crops. With species extinction rates at an all time high (Dirzo et al. 2022; Munstermann et al. 2022), much of the uncatalogued biodiversity could be lost before it is understood scientifically. Participation gaps likely contribute to taxonomic sampling gaps in complex and context dependent ways. For example, the exclusion of Indigenous perspectives from science has removed valuable knowledge of local biodiversity and diverted resources away from regionally important plants (Dwyer et al. 2022). Together, these factors exacerbate the patriarchal and eurocentric system of publication, and result in a body of literature that poorly represents the global wealth of biological diversity and knowledge.

To better understand the changing global landscape of plant science research and quantify patterns of underrepresentation, we conducted a comprehensive bibliometric analysis of nearly 300,000 papers published across the past two decades of plant science research. Our analyses are framed from the perspective of the first axiom of Ardila-Mantilla which states that scientific potential is “distributed equally among different groups, irrespective of geographic, demographic, and economic boundaries” (Ardila-Mantilla, 2016). If we take such a statement to be the null hypothesis, then disparities in educational advancement and promotions, funding, or publication and citation rates indicate that other factors, like oppression, have created historical and contemporary biases in science. To test this hypothesis, we begin by identifying demographic features that are associated with high publication and citation rates in plant science. We then quantify taxonomic sampling gaps and regional differences in focal organism choice to explore associations between participation gaps and study organisms. We examine how these dynamics change over time and space to identify areas that are improving, stagnant, or regressive. We close by discussing the need to dismantle oppressive systems in the plant sciences, improve equity, and how such changes will ultimately advance the field in the coming decades.

RESULTS

We compiled a comprehensive database of 296,447 plant science papers published between 2000 to 2021. Papers were sourced from a representative set of 127 plant science journals based in 26 different nations across 5 continents, covering 21 different subspecialties. We included both society and for-profit journals in our analyses, with open access, hybrid, and subscription publishing models (see supplementary Table S1 for journal information).

Geographic disparities in publication and citation rates

To gain insight into the global landscape of plant science research, we summarized geographic differences in publication and citation numbers. We predicted that researchers working in low-income nations, historically excluded communities, and under-resourced institutions would face multiple barriers to publishing in high impact, highly cited journals, including reduced access to specific technologies, exclusion from collaboration networks, and financial constraints.

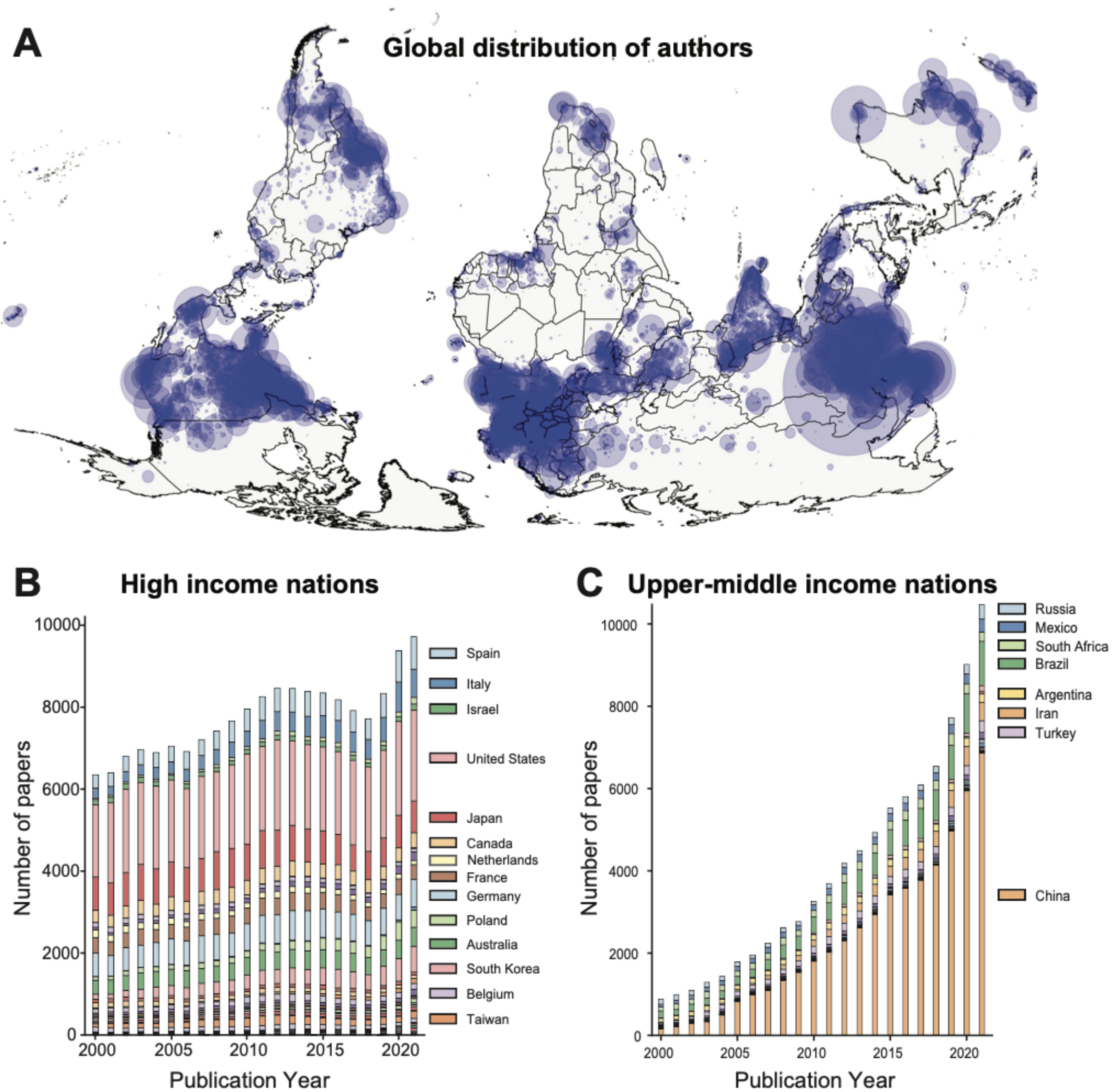

Figure 1. Global patterns of plant science publishing. A) The global distribution of where authors are based, scaled by the number of publications from each location. B) The number of studies published each year by authors in high income nations. C) The number of studies published each year by authors in upper-middle income nations.

We found strong evidence to support this hypothesis. Vast areas of the world have virtually no footprint in the plant science literature over the past two decades (Figure 1A) and publication rate is tightly correlated with national affluence. On a continental level (Supplementary Figure 1A), nearly one third (27%) of all papers were led by authors based in Europe, another 18% were led by authors in Northern America, and 37% by authors in Asia. The remaining 17% of publications were led by authors distributed across Africa, Latin America, and Oceania. Within each continent, authors were further consolidated into distinct hubs of research activity, with the USA, China, and Western Europe dominating the plant science landscape (Figure 1A). Patterns related to national income were extremely skewed. Based on the United Nations's income classifications of high, upper-middle, lower-middle, and low income (Supplementary Figure 1B), we found that 61% of all papers published in the last 20 years were led by authors in high income nations. Another 32% were led by authors in upper-middle income nations, and the remaining ~7% of publications were distributed among lower-middle nations. Less than 1% of papers were led by authors in low income nations. National publication rates were highly correlated with Gross Domestic Product (GDP) (R^2=0.96, $F_{1,141}$=3404, p=1.28e-100) (Figure 2B), but there was very little correlation between publication rate and per capita income (R^2=0.073, $F_{2,136}$=5.39, p=0.0056) (Figure 2C). Interestingly, research output in high income countries has remained relatively stable over the past 20 years, with most nations showing only modest increases in publications over the study period. In some high income nations (e.g. the Netherlands), research output actually declined over the past 20 years. In contrast, there has been a 10-fold increase in the number of papers from upper-middle income nations in the past two decades. In fact, by 2021, there were more papers published by authors in upper-middle income nations than by authors in high income nations (Figure 1B and C). However, this increase was driven primarily by China, which accounted for more than 60% of the publication output from upper-middle income nations in 2020. Other emerging economies such as India, Brazil, Iran, South Africa, Mexico, and Argentina have also made noteworthy contributions to the increased research output of upper-middle income nations (Figure 1C). Publication rates in lower-middle and low income nations have also increased in the past two decades, but still lag far behind those of high and upper-middle income nations (Supplementary Figure 2). In some cases, noticeable decreases in research activity appear to correlate with national disasters and war (e.g. Syria's annual publications declined sharply in the past 10 years (Supplementary Figure 2A)). Despite noteworthy growth in plant science research, many countries remain underrepresented in the literature.

In general, productivity is expected to scale with population size following a power law,w such that larger cities produce more research output than smaller ones (Bettencourt et al. 2007a; Bettencourt et al. 2007b; Bettencourt et al. 2010). In plant science research, we observed a rough power law (reflected by a linear relationship in logarithmic plots) of papers produced versus population size (Supplementary Figure 3). However, this scaling was variable across the globe. In general, cities in Northern America, Northern Europe, and Oceania had above average research output relative to population size. In contrast, cities in Asia, Africa, and Latin America had below average research output relative to their population size (Figure 2A). Taken together we find that high income nations produced a higher proportion of their research in rural areas, whereas lower income nations concentrated research activity in high-density, urban areas. This is noteworthy because high income nations in Northern America, Europe and Oceania account for less than 10% of the rural population globally (Supplementary Figure 4), but produce more than 64% of the plant science research.

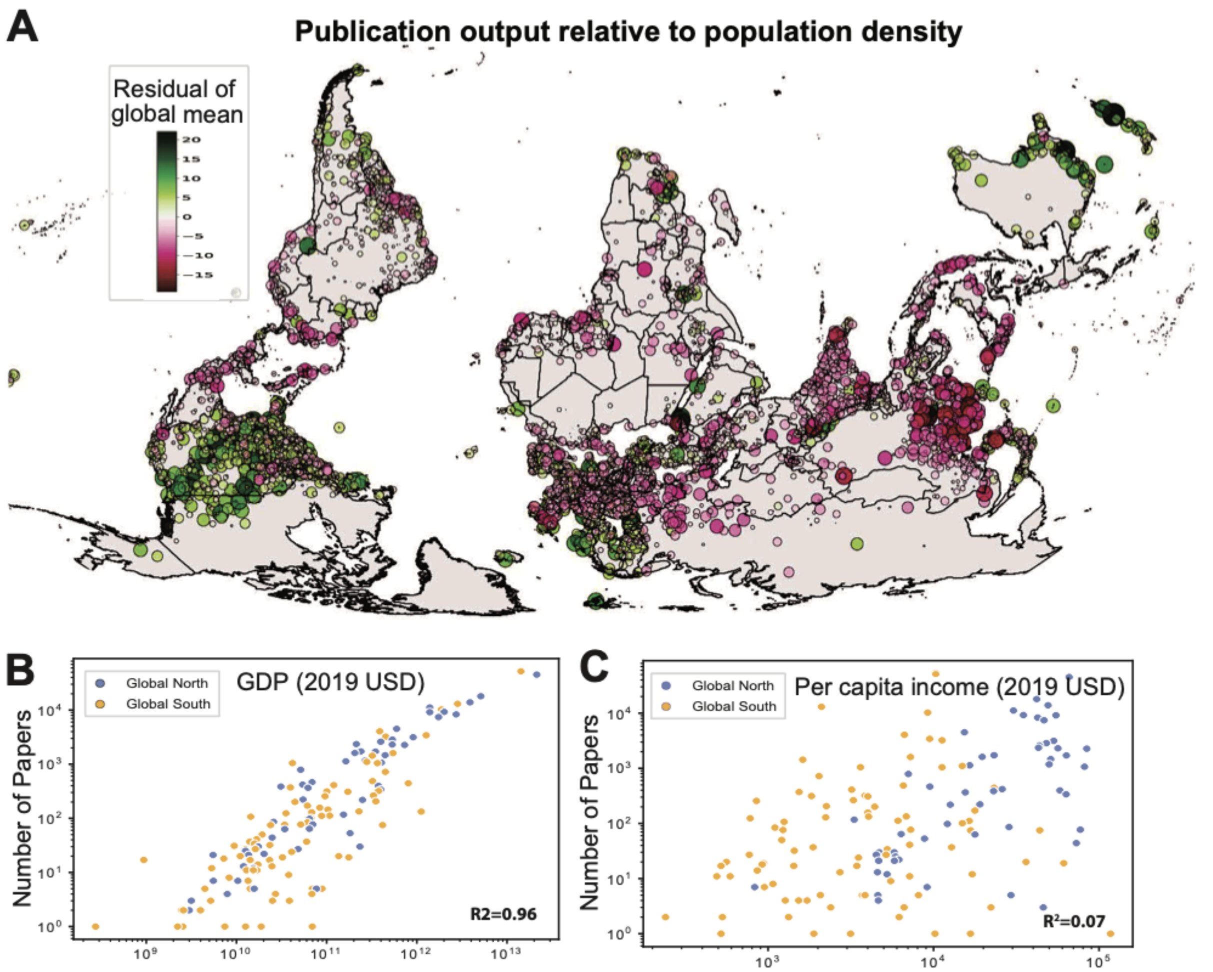

Figure 2. Publication output relative to population size and national affluence. A) Map of publication output relative to population size for locations with more than 300,000 inhabitants or with more than 100 papers produced during the time period of 2000-2021. Locations are scaled and colored according to their research output relative to the global trend. Large green points correspond to locations that produce more research than expected based on the global population trend, while large pink circles represent regions that publish less than expected for a city of their size. B) The number of studies published by each nation relative to national GDP. C) The number of studies published by each nation relative to per capita income.

International and intercontinental collaborations were strikingly uncommon in the past two decades of plant science research (Figure 3 and Supplementary Figure 5). More than two thirds (71%) of the publications in our database were written by authors based in a single nation. Just 22% of studies involved a collaboration between two nations, and only 5% of studies included three nations. Only 1% of studies involved four nations even though 71% of papers have four or more authors, and just 0.04% included five nations despite the fact that 54% of papers had five or more authors. When international collaborations did occur, they tended to be across continents rather than within continents. Only Europe-based authors showed a high frequency of within-continent collaboration (Supplementary Figure 5). Collaborations across continents did occur, but were not evenly distributed. Most nations preferred to collaborate with researchers in Europe, Northern America, or China (Figure 3), and were less likely to collaborate with authors in Latin America, Africa, or West Asia. A similar pattern is evident when considering income groupings–only the most affluent nations participated in within-group collaborations, and all other nations preferred to collaborate with high income nations (Supplementary Figure 5).

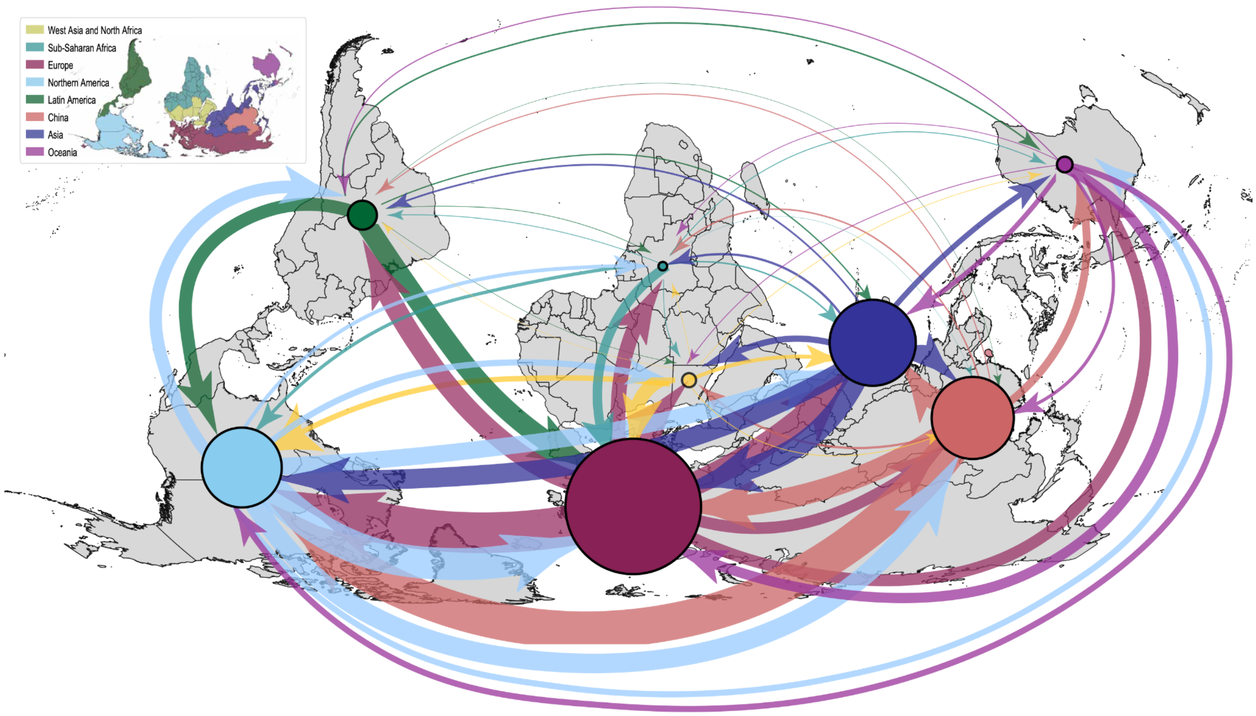

Figure 3. Disparities in global collaborations within plant science research. Circles represent publications that did not involve an intercontinental collaboration. Arrows represent cross-continental collaborations and are directed from corresponding author to co-author. Circles and arrows are scaled by the number of publications.

Despite striking differences in research output across continents, the mean impact factor of publications spanned just over one point across continents–ranging from 2.92 ± 0.017 for papers led by authors in Sub-Saharan Africa to 4.06 ± 0.011 in Northern America (Table 1). By contrast, citation rates were substantially more variable across continents. In general, papers from the Global South received dramatically fewer citations than those from the Global North, despite publishing in journals with similar impact factors. For example, mean cumulative citations ranged from 17.82 ± 0.304 for papers led by authors working in Sub-Saharan Africa to 36.75 ± 0.298 in Northern America (Table 1). This dynamic has remained relatively stable over the past 20 years, with persistent differences in annual citation rates between continents (Supplementary Figure 6). Some individual nations (e.g. China) have seen improvements in citation rates over time, but most have not.

Table 1.The mean impact factor of journals that authors published in and mean number of citations that papers received for authors from different continents.

Continent	***Mean impact factor***	***Mean cumulative citations***
Northern America	*4.06 ± 0.011*	*36.75 ± 0.298*

Oceania	*3.67 ± 0.022*	*31.99 ± 0.621*
Europe	*3.94 ± 0.008*	*31.21 ± 0.193*
Asia (minus China and West Asia)	*3.53 ± 0.008*	*26.32 ± 0.210*
North Africa and West Asia	*3.05 ± 0.017*	*23.00 ± 0.409*
China	*4.13 ± 0.009*	*21.69 ± 0.159*
Latin America and the Caribbean	*3.13 ± 0.011*	*18.57± 0.240*
Sub-Saharan Africa	*2.92 ± 0.017*	*17.82 ± 0.304*

We also investigated how journal policies such as open access fee, impact factor, and society membership related to participation rates for authors with different identities. We predicted that author diversity would decrease with increasing open access fees, that “elite” high impact factor journals would exhibit reduced diversity, and that society journals would be more inclusive. Of the 296,447 papers examined, only 14% were published Gold open access. Authors in Northern America and Asia published the highest proportion of open access papers (23% and 18% respectively). In contrast, only 10-15% of papers led by authors based in Africa, Latin America, and West Asia were published open access. Of the 16,641 papers published in “elite” journals with impact factors above seven, 68% were led by authors in high income nations, compared to 61% overall, and another 15% were led by authors based in China. The remaining 17% were distributed across authors in lower income nations. Citation rates were also extremely skewed within these journals. For example, papers led by authors in high income nations were cited 82 ± 0.23 times whereas papers from low income nations were cited only 24 ± 0.86 times–a four fold difference. In general, society journals did not exhibit any more geographic equality in publication and citation rates than the overall trend. Of the 158,711 papers published in society journals 63% were led by authors in high income nations and these received almost double the number of citations (38.9 ± 0.217) compared to papers led by authors in low income nations (20.3 ± 1.767).

Persistent gender inequalities in the plant sciences

We quantified the effects of patriarchy and gender discrimination in plant science publishing by associating author names with masculinity or femininity. We acknowledge that a binary gender division is an oppressive concept in itself and that true gender is self-identified (Castro-Peraza et al. 2019), and we recognize the proximity of our approach to the harmful practice of gender inference. However, we find that we cannot discuss patriarchy and gender discrimination without employing the concept of gender. We purposefully seek to avoid inferring the gender identity of individuals and rather measure the oppressive effects of patriarchy associated with the names themselves. We focus on the normative association of names with masculinity or femininity to measure these effects, and do not presume to know the true gender identity of authors. We struggled with the ethics of algorithmic gender inference within our working group and must acknowledge that in conducting such analyses, we too are culpable in propagating the gender binary anew (Martin, 2019). We further acknowledge that biases in name-based gender inference can arise from the global diversity of cultural naming systems (Santamaría and Mihaljević, 2018).

We hypothesized that individuals with marginalized gender identities (including women, non-binary, gender non-conforming, trans, and people of multiple sexes/genders) would face barriers to participation in plant science and that these would compound with socioeconomic disadvantages and/or historical oppression to further limit participation by intersectional individuals. We cannot test this hypothesis directly without knowing the gender identities of the authors in the paper, so we aimed to instead measure perceptive discrimination based on sexism, that disadvantages individuals with names normatively associated with femininity (Topaz et al. 2022; Maril & Gill 2018). To test this prediction, names of corresponding authors for each paper were classified as either 1) names normatively associated with masculinity (NNMs) or 2) names normatively associated with femininity (NNFs) and used as a proxy for gender.

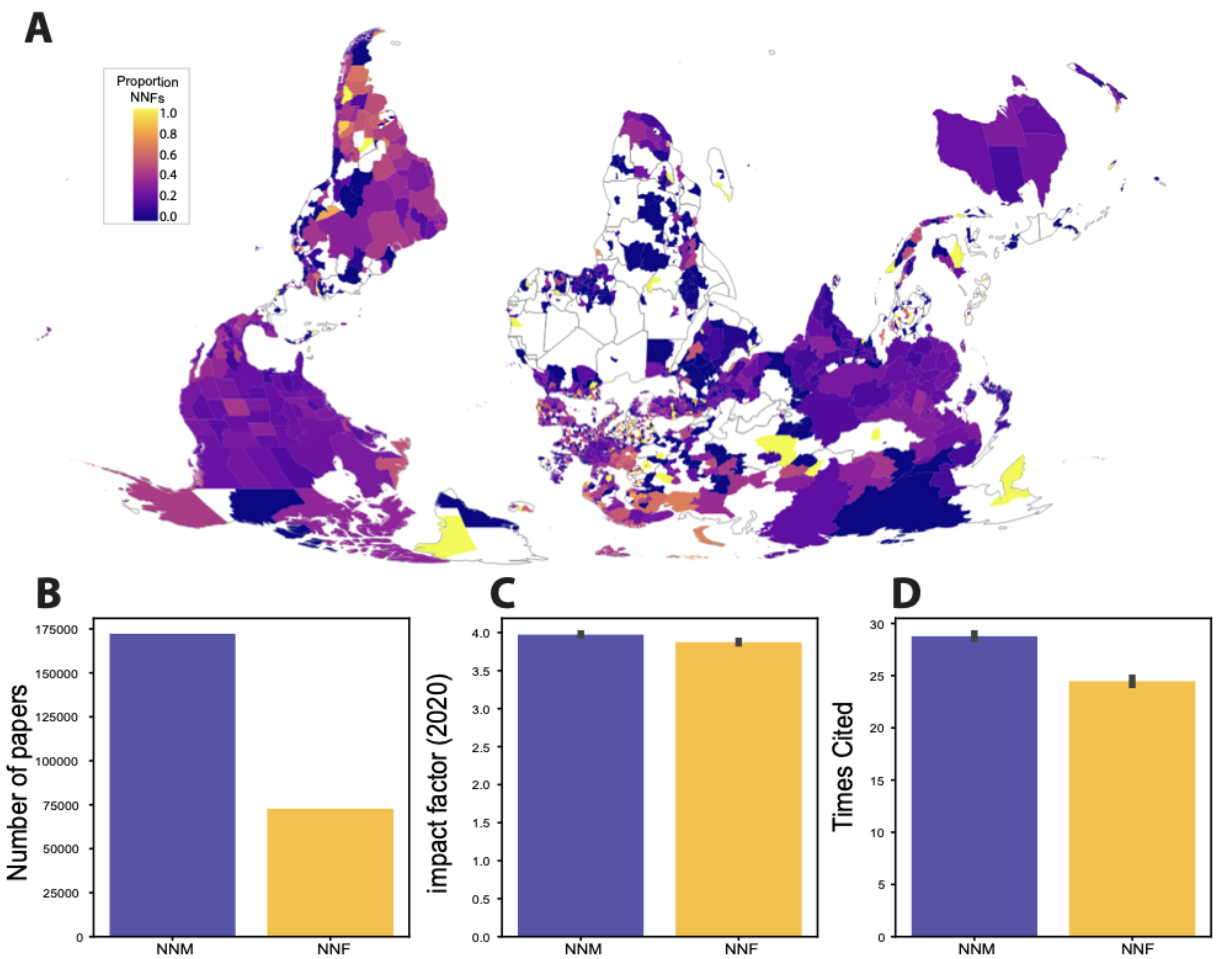

Figure 4. Disparities in the global distribution of corresponding authors by gender. A) Map showing the distribution of the gender ratio of names. There were many more papers led by authors who had names normatively associated with masculinity (NNMs) than papers led by authors with names normatively associated with femininity (NNFs), but the extent of the imbalance was highly variable across

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.15.512190; this version posted October 18, 2022.

the globe. B) The total number of publications led by authors with NNMs and NNFs, C) the impact factor of the journals published in, and D) the citation rates for papers led by authors with NNMs and NNFs.

Globally, there were far more papers led by authors with NNMs than authors with NNFs (Figure 4A and B). However, the degree of gender imbalance was highly variable across continents and nations. Among the 20 nations with the highest publication rates, the most NNM baised nations were Japan (14% NNF), India (21% NNF), Netherlands (23% NNF), Switzerland (24% NNF), and Israel (25% NNF). In contrast, the least NNM biased nations were Poland (61% NNF), Argentina (57% NNF), Italy (41% NNF), Brazil (41% NNF), and Spain (38% NNF). On a continental level, Latin America and Europe had the highest proportions of papers led by authors with NNFs whereas Northern America, Asia, and Oceania had the lowest proportion of NNFs. There has been a modest increase in participation by individuals with NNFs over time, but gender ratios remain far from equal across much of the globe (Figure 5).

There was no correlation between national GDP and the proportion of papers led by NNFs (R^2=-2.3, $F_{1,117}$=-81.9, p=1) (Supplementary Figure 7A). In fact, some of the highest GDP nations had the lowest proportion of NNF authors. However, there was a weak positive relationship between per capita income and the proportion of NNF authors (R^2=0.139, $F_{2,113}$=9.1, p=0.0002) (Supplementary Figure 7B).

There was no significant difference in the impact factor of the journals that authors with NNFs versus NNMs published in. However, there were noteworthy differences in the number of citations these papers received. Papers led by authors with NNMs were cited on average 5 more times than those led by authors with NNFs. This pattern has not improved over time and, if anything, the difference in annual citations for authors with NNFs versus NNMs has expanded (Supplementary Figure 8).

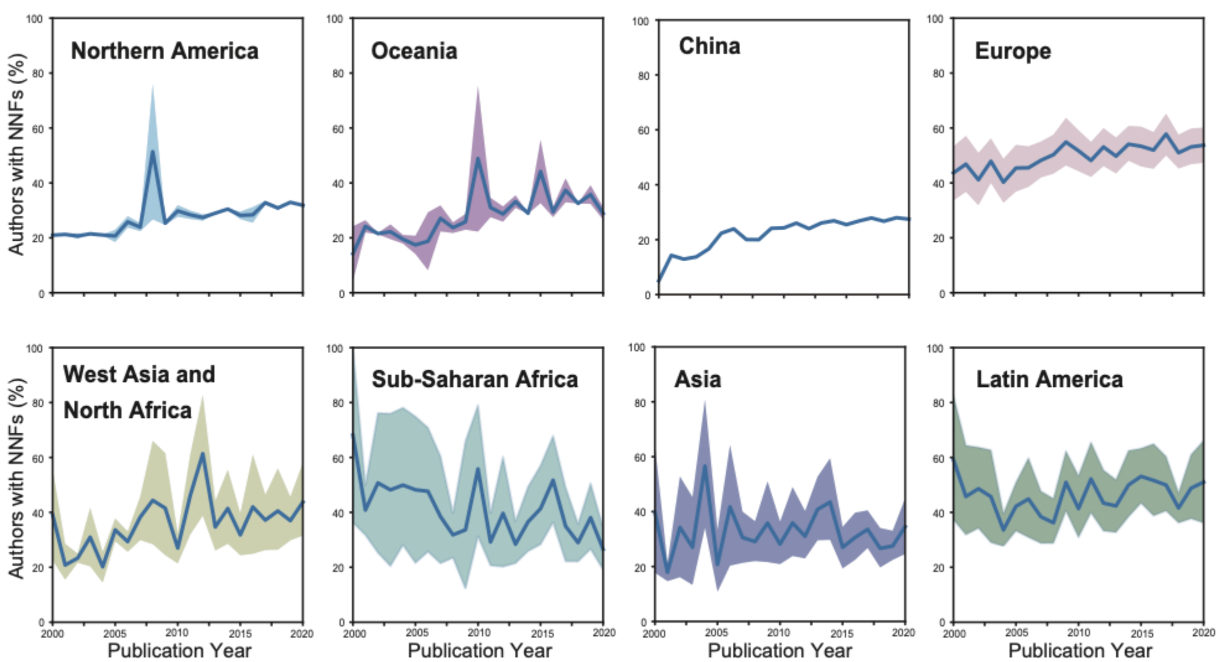

Figure 5. Stagnant gender bias over the past two decades. The proportion of authors with NNFs over the last 20 years is plotted for each of the eight geographical regions investigated. Solid lines show the mean and shading represents 95% confidence limits.

Taxonomic gaps in focal species studied in the plant sciences

Funding priorities and research activities have historically focused on a narrow subset of described plant species (Trimble and van Aarde 2012; Adamo et al. 2021; Marks et al. 2021) and we expected to find notable taxonomic sampling gaps in the plant science literature. To test this prediction, we identified all taxonomic entities mentioned in abstracts via natural language processing. We then summarized overall patterns and geographic differences in the choice of focal species to identify taxonomic sampling gaps and regional patterns.

There were 73,527 unique taxonomic entities represented in our publication database. While the majority of studies focused on plants, we also identified numerous non-plant species including pathogens, symbionts, and other interactors across animalia, fungi, and bacterial groups (Figure 6B). All of the top 20 most studied plants represent economically important crop species or models developed by the plant research community (Figure 6A). The model plant *Arabidopsis thaliana* was by far the most studied plant in the past two decades, appearing in four times as many studies as the next most common species wheat (*Triticum aestivum*) (Figure 6A). Poales was the most studied order with over 50,000 mentions, followed by Brasicales, Fabales, and Solanales (Supplementary Figure 9). Many orders were statistically over- or under-represented in the dataset relative to their species richness. The most over-represented orders were Brasicales, Poales, Solanales, Fabales, and Cucurbitales. In contrast, the most under-represented clades were Asterales, Asparagales, Gentianales. Polypodiales, and Lamiales (Figure 6C).

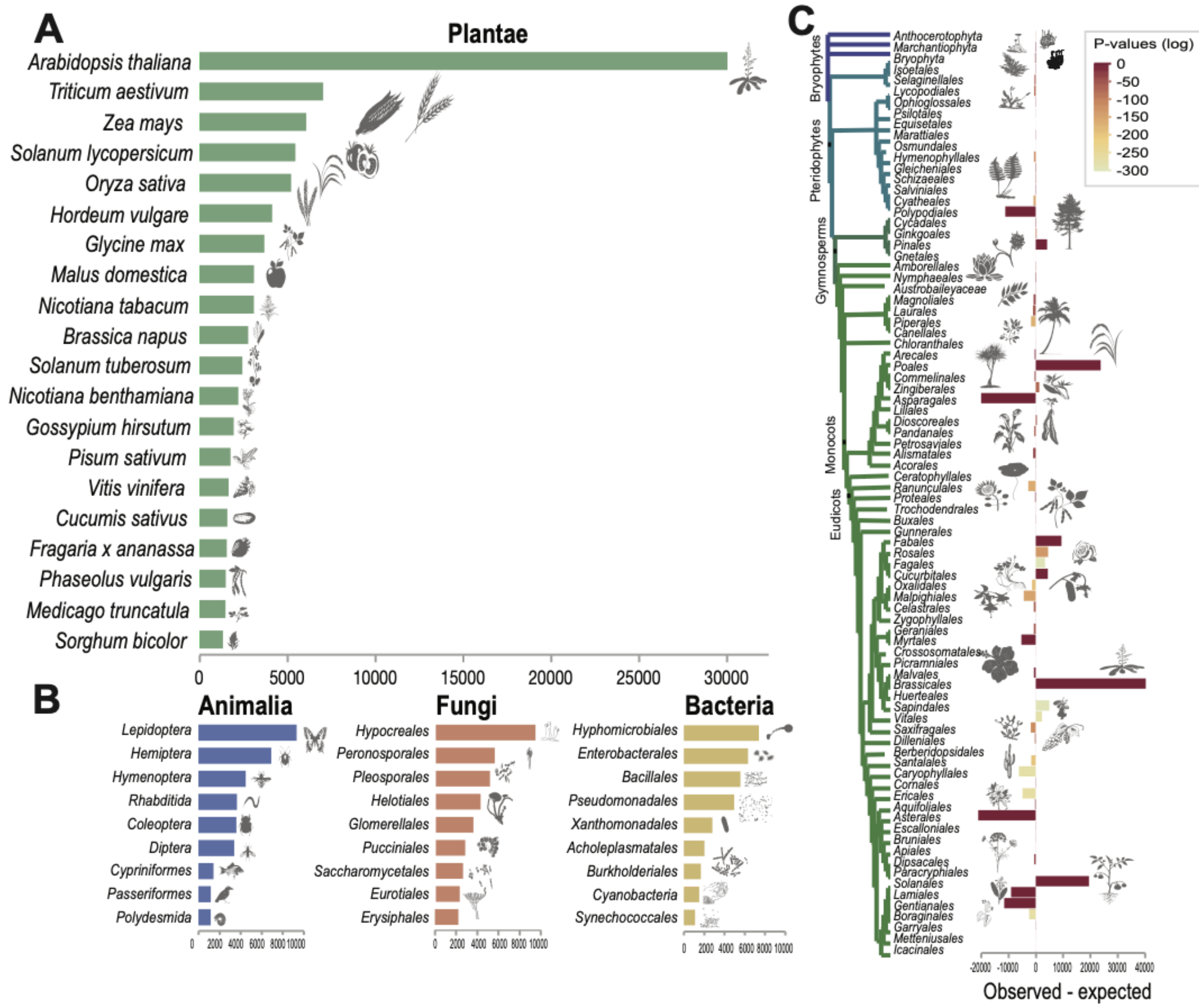

Figure 6. A) The top 20 most studied plant species across all studies. B) The top 9 most studied orders for non-plant groups (animalia, fungi, and bacteria). C) The observed number of studies investigating each order or land plants minus the number expected if sampling effort had been evenly distributed relative to species richness.

We also identified regional differences in the choice of focal organisms. Most high income nations with high publication rates tended to focus on *A. thaliana*, grain crops, vegetables, fruits, and model species (Figure 7). In contrast, many of the nations underrepresented in publishing, tended to focus on lesser known species and minor or regionally important crops. This finding exemplifies how underrepresentation at the human level impacts the diversity and breadth of focal organisms and research directions.

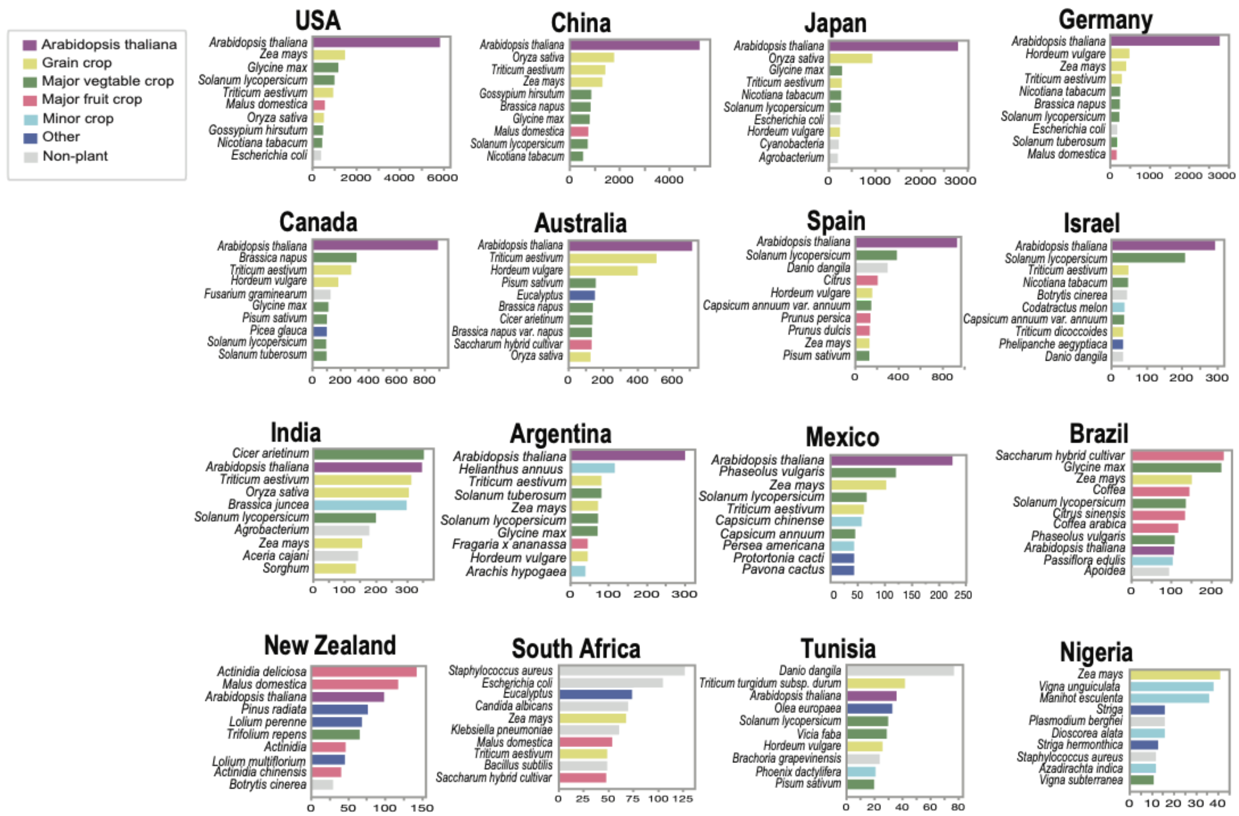

Figure 7. National differences in focal organism choice. The top 10 most studied species from the literature in select nations is plotted. Nations are organized from most prolific to least, and focal organisms are colored by generalized groupings of organism type. The x-axis shows the number of papers that focus on each focal organism.

DISCUSSION

Our analyses reveal striking geographical biases in plant science research that are associated with national wealth. We identified a strong correlation between publication rates and GDP but only a weak association with per capita income, suggesting that research activity is driven by investment at governmental and institutional levels rather than by individuals. Global patterns of wealth distribution cannot be understood without an acknowledgement of the impact of imperial colonialism and the resulting consolidation of resources within select nations of the Global North (Adas 2008; Baber 2016; Marks et al. 2021). Not only was wealth redistributed during this process, but diverse perspectives and peoples were effectively erased as a eurocentric worldview was exported across the globe. Christian missionaries, European traders, and inquisitive researchers all helped to spread frameworks of capitalism, patriarchy, and white supremacy. In biology, these value systems are coupled with a precedence for the English language, reductionist thinking, Latin naming conventions, and biased standards of academic excellence that further exclude individuals from non-European backgrounds.

We found that research output is also associated with increased population density, where more papers are produced in more populous areas. However, we detected regional differences in this pattern. High income nations (especially in Northern America and Oceania), generated a substantial proportion of their research in rural areas. This makes intuitive sense since plant science is inherently linked to agriculture and natural spaces, and numerous research centers and Land-grant Universities have been built in rural regions. However, lower income nations did not produce many papers in rural areas and instead concentrated research activity in urban centers. We suspect that this pattern is driven by the fact that rural areas are often the last places to be developed and still lack basic infrastructure across much of the globe (Lagakos 2020). These differences in rural development impact where research is conducted and contribute to the exclusion of rural peoples and agricultural communities in less affluent nations. Only 8% of the world's rural population lives in Europe, Northern America and Oceania, but these areas produce more than 64% of the plant science papers. The remaining nations have a disproportionately small publication footprint and the knowledge of ecology, ethnobotany, and agriculture from local and idiginous communities within these countries is largely absent from the literature. These voices and perspectives are often co-opted by Western researchers through parachute science and other colonialist practices, with no acknowledgment, consultation, or compensation for the discoveries (Stefanoudis et al. 2021). These harmful practices have perpetuated persistent inequity in the field. Such gaps in participation undoubtedly translate into gaps in understanding and represent a lost opportunity.

International and intercontinental collaborations were notably uncommon in the past two decades of plant science research. Of the few international collaborations that we identified, the majority involved a collaborator from Europe, Northern America, and to a lesser degree, China. We suspect that differences in resources (both financial and infrastructural) contribute to these dynamics. Researchers working in high income nations have access to more funding for international research, engaging collaborators, and traveling to conferences. Researchers in less affluent nations do not have the same funding opportunities and are therefore limited in the number and type of collaborations they can participate in. There may also be more subtle and problematic factors driving the skewed collaborative networks we observed. Differences in institutional prestige born out of eurocentric mindsets, have led some to believe that the best science is done in select institutions in the Global North and that working at or collaborating with those institutions is most desirable. We believe that this rationale is fundamentally flawed and should be dismantled. Affluent nations could do more to engage collaborators in less represented regions of the globe instead of following the well-established global network. Not only would this help to equalize the plant science landscape, but it would enrich our science by bringing in the wisdom of different perspectives.

We identified striking and persistent gender biases in plant science publishing. Over 70% of publications in the past two decades were led by authors with masculine names. The extent of gender imbalance was variable across nations and continents, but showed remarkably little change over time. In most regions, we detected only modest increases in the number of papers led by authors with feminine names over the past two decades. Interestingly, some of the most affluent nations (e.g., USA, Japan, Netherlands, Switzerland, Germany, Canada, and New Zealand) had extremely male biased publication records despite supposed improvements in women's rights in the Global North. In contrast, some less affluent nations in the Global South (e.g., Argentina, Brazil, and Mexico) had some of the highest proportions of lead authors with NNFs. This finding is similar to the "gender-equity paradox" detected in mathematics (Breda et al. 2020), and contradicts our prediction that individuals facing the intersecting barriers of economic constraints and marginalized gender identity would be more excluded from academic publishing. It suggests that other factors, like cultural differences, could be playing a role in gender inequity. For example, in regions where farming and agriculture are traditionally women's work, more women may choose to enter the plant sciences. In addition, differences in available support systems can drive career choice, with women sometimes pursuing higher paying jobs (often in STEM fields) when social support systems are limited (Stoet and Geary 2018). We looked at a variety of economic development indicators to try to understand what could be driving gender biases in plant science publishing. In contrast to geographical patterns, there was no association between national GDP and gender ratio. There was however, a slight association between per capita income and the proportion of authors with NNFs, suggesting that participation by authors with NNFs in plant science could be related to individual economic mobility rather than national investment. Regardless of the reason, these findings suggest that the footprint of patriarchy in plant science is deeper than we acknowledge and does not align neatly with narratives about cultural differences in sexism. We also identified gender biases in citation rates that were independent of time, suggesting persistent and ongoing gender descrimination. Because individuals, not institutions drive citation rates, this suggests a deep and pervasive bais running through the discipline. It also means that we, as individuals, have the power to shift these patterns through our actions and choices.

In the past two decades, plant scientists have studied thousands of species spanning plants, animals, bacteria, and fungi. Despite the noteworthy diversity and volume of research, sampling effort has not been equally distributed across clades and taxa. The vast majority of studies have investigated major crop and model species, and the remaining biodiversity accounts for only a fraction of the research on plants. Our analyses identified a number of statistically overrepresented groups of plants, all of which included agriculturally and economically important plants. We also identified numerous underrepresented taxonomic groups, which were ecologically diverse, speciose, and generally of less economic relevance to modern society. These underexplored lineages could provide untold value to humans and ecosystems, but have been largely overlooked by modern plant scientists (Hendre et al. 2019; Marks et al. 2021; Dwyer et al. 2022). We found some evidence to indicate that taxonomic gaps are related to geographic and gender gaps and that limited diversity of authors is exacerbating biases in study organism choice. In general, affluent nations in Europe, Northern America, and Asia tended to focus on major crops associated with industrialized agriculture (e.g., wheat, rice, soybean, tobacco, tomato, etc.). In comparison, many of the nations with a smaller footprint in plant science focused their research on regionally important and underutilized crops such as cassava, yam, and millets or local plants with medicinal or historical importance. The disproportionate focus on major crops reinforces the homogenization of agriculture and limits our understanding of and ability to conserve and utilize biodiverse plants. We suspect that if more researchers from across the world were actively engaged in plant science research, there would be a diversification of study systems and a broadening of cumulative knowledge.

Conclusions

Our analyses provide evidence of deep disparities in plant science with links to colonialism, eurocentrism, and patriarchy. Despite the proliferation of statements, committees, workshops, and trainings aimed at increasing diversity, equity, and inclusion, little progress has been made towards actually diversifying plant science in the past two decades (Montgomery and Whittaker, 2022). Our results suggest that a eurocentric value system permeates scientific research and continues to exclude and marginalize some individuals. These comprehensive findings can be used as motivation and evidence in advocating for change at institutional and policy levels. While many recognize that the current system is unfair, there are contrasting views on what changes should be made. Some advocate for reformation while others favor abolition, but both agree that there is a need to broaden science and embrace the diversity of knowledge acquisition systems that exist globally. We suggest that first steps towards improving the discipline should consist of a fundamental broadening of our definition of what science is and who can do it. By embracing a more nuanced and context dependent view of data, acknowledging that novelty is not the only source of scientific merit, and recognizing the value of qualitative research, we can begin to minimize colonial biases in academic culture, language, and institutions (Dwyer et al. 2022). Funding is another important component, and wealthy nations should take the lead in making efforts to equalize disparities in national affluence established through colonialism. Grants that specifically promote intercontinental collaborations could play an important role coupled with direct funding to lower income nations. Given the longstanding disparities that exist in plant science, it may be useful to employ concepts of restorative justice including truth and reconciliations practices (Wong et al. 2020; Montgomery and Whittaker 2022), and a more general shift away from gatekeeping policies and towards inclusive concepts like groundskeeping that attend to the needs of individuals while also supporting the health of institutional ecosystems (Montgomery 2020). By expanding our definition of what constitutes scientific inquiry and who can take part in it, we open the door to new sources of knowledge. Rather than continuing to convert people to eurocentric ideals and ways of knowing, we should make space for other systems of knowledge to rise to the forefront. We hope our comprehensive set of analyses can be used to support these positive changes.

METHODS

Data acquisition and filtering

We assembled a comprehensive database of plant science papers from 127 journals spanning a range of impact factors, nationalities, and sub-specialties (see supplementary table S1). We began with a list of plant science journals put together by the American Society of Plant Biology (https://plantae.org/plant-biology-journal-database/), which we filtered on the following criteria: (1) the journal must have an impact factor, (2) it must be plant specific, and (3) it must include research articles. We then added another 53 additional journals based on a search for “plant sciences” journals on the Journal Citation Reports Database (https://jcr.clarivate.com). We filtered these journals using the same filtering criteria as previously mentioned. Research papers from the resulting 127 journals were included in the current study.

Metadata for all research articles published in these journals during the years 2000-2021 were downloaded from the Web of Science (WoS) using the batch download tool. For each paper the downloaded metadata included the *Author Full Names, Article Title, Author Keywords, Keywords Plus, Abstract, Addresses--all authors, Address--corresponding author, DOI, Email*

Addresses, Researcher Ids, ORCIDs, Funding Orgs, Funding Text, Cited Reference Count, Times Cited All Databases, ISSN, eISSN, Open Access Designations, Quartile, Publication Year, Volume, Issue, IDS Number, UT (Unique WOS ID), and *Pubmed Id*. The resulting dataset was filtered to remove duplicate records, papers without a corresponding author, all book chapters, reviews, proceeding papers, and retracted papers. A total of 296,447 records were retained across all 127 journals. The complete dataset, along with a description of data acquisition and curation are deposited in Dryad at https://doi.org/10.5061/dryad.pg4f4qrtb.

Other metadata were incorporated by referencing JCR and journal webpages, including *Journal Impact factor (2020), Publisher Publisher City Publisher Address, Journal location, Subspecality (JCR), Open Access options available, Open access fees (USD)*. We consolidated open access designations into two categories for simplicity. Papers were scored as open access if they were published Gold open access. All other open access designations (e.g., green, bronze, etc.) were not considered open access. For more information about open access designations please see (https://clarivate.com/blog/a-researchers-complete-guide-to-open-access-papers/). Data on national development indicators were taken from the World Bank 2019 database (https://databank.worldbank.org/source/world-development-indicators). Continental divisions and country assignments were based on a sensible combination of subregions as designated by the UN Statistics Division (https://unstats.un.org/unsd/methodology/m49/). Income divisions were based on per capita GNI in June 2019 as reported by the UN Department of Economic and Social Affairs (World Economic Situation and Prospects).

Geography based analyses

The location of authors was inferred from the addresses listed in the papers. Using an ad-hoc text processing script, the location of every author was defined as the city, regional administrative division, and country as extracted from the address. We kept track of addresses associated with the corresponding authors vs addresses associated with the rest of co-author locations. We then elaborated a list of unique locations and tallied the number of repetitions for each. We were unable to either link individual authors to individual addresses, compute how many authors were associated with each address, or keep track of authors with multiple addresses and affiliations. We were thus limited to consider every individual address as one separate paper for global tally purposes.

Geographic coordinates (geocoordinates) for all these locations were obtained using the Google Maps Geocoding API (https://developers.google.com/maps/documentation/geocoding) with Python via GeoPy. Geocoding API failed to provide geocoordinates for a small number of locations (24), which we extracted manually. Locations that were within 25km from each other were merged and their tallies combined to account for extended metropolitan areas and name changes. For example, the papers associated with Coyoacán, Astana, and Hyderabad, Andhra Pradesh were merged with those from Mexico City, Nur-Sultan, and Hyderabad, Telangana respectively. These merges were performed only when both locations were within the same country to avoid merging cities located at national borders. Ultimately this pipeline produced 7,121 unique locations spread across 201 countries and territories.

Each location was associated with the population of the closest city listed in the GeoNames database (https://www.geonames.org/). We also added all the population of any other city within 25km to account for a broader metropolitan area. As in the previous step, these merges were limited to cities within the same country. Additional entries and census data were manually added to our local copy of the GeoNames database to account for locations that failed to have a city and population assigned. A global power law of papers published with respect to population size was computed by first log-transforming all the research output and population size data, and then fitting a reduced major axis (RMA) linear regression (Smith 2009). RMA was chosen to account for the variability of the population size, the x-axis, as it varied over the period when the papers were produced. Residual deviate products were computed as the true residuals for RMA fitting (Uylings et al. 1986). These residuals were interpreted as the scale-adjusted metropolitan indicators (SAMIs), as proposed by (Bettencourt et al. 2010), to identify over-performing and under-performing plant science research locations given their population size. To account better for geographical differences, separate scaling models were computed and compared by just using locations related to countries in a specific continent or income bracket.

For every case, especially for countries, we respected the location provided by WoS data. We however recognize that the full addresses listed by WoS might not fully match the addresses listed in the original papers. For example, all institutions based in Hong Kong and Puerto Rico have addresses that list them as part of China and the United States respectively. Similarly, we respected WoS designation for locations under territorial dispute. For example, Sevastopol was listed as part of Ukraine for some addresses, while it was listed as part of Russia in others. We also respected WoS designation of certain territories as individual countries despite their lack of worldwide recognition, such as Palestine, Kosovo, and French Guiana. Addresses corresponding to countries and territories that are no longer present were manually examined. All addresses listed as part of Yugoslavia or Serbia and Montenegro were manually assigned to Serbia, as the institutions listed are physically located within modern day Serbia. Similarly, all addresses from the Netherlands Antilles were assigned to Curaçao. Finally, country names were updated to reflect their most recent name, as in Czechia, eSwatini, North Macedonia, and Türkiye.

To quantify patterns of collaboration, we identified the corresponding author for each paper and assigned the paper to that location. We then assessed the locations of all other co-authors on that paper and determined if they were from different countries, continents, or income brackets. Summary stats were computed using Python packages Pandas and Numpy and visualized in Seaborn, Cartopy, and Matplotlib. The supplementary alluvial plots were done in R with the ggplot2 package.

Gender analyses

The analyses presented here do not identify the true gender of authors. Rather, they show the assumed gender based on the association of first name with either masculinity or femininity. These analyses also likely mis-identify and fail to account for non binary, gender neutral, and trans individuals, among others. Geographic biases in the performance of gender inference have also been documented, with most tools performing poorly on east Asian names. This is noteworthy since many of the papers in our dataset are led by individuals with east Asian heritage. Given these caveats, we selected the most robust tool (GenderAPI) available for this type of analysis based on the extensive benchmarking and comparative analyses presented in (Santamaría and Mihaljević 2018).

To identify the first names of authors, we used natural language processing to extract the full name of the corresponding author from the complete list of author names. This yielded complete first names for ~60% of the papers in our dataset. The remaining 40% of entries had various forms of abbreviated names. We manually curated the resulting data to fill in gaps and consolidate abbreviated names. We used the following rules to obtain full names from abbreviated forms and initials: (1) authors with matching last name and first two initials were considered the same author and given the longest form of the first name; (2) authors with matching last name, first initial, and email were considered the same author and given the longest form of the name; and (3) authors with matching last name, first initial and institution, or matching first initial and unique last name were considered the same author and given the longest form of the name. If the author identity was ambiguous (e.g., two authors with a shared last name and first initial) we left the field blank. Manual curation allowed us to recover the first names for an additional 25% of the records in our dataset. We then submitted these first names along with the nationality to genderAPI to infer the gender of author names. Summary stats, national gender ratios, and changes over time were computed using Python packages Pandas and Numpy and visualized in Plotly, Seaborn, and Matplotlib.

Study Species analyses

The species studied in each paper were identified from abstracts using the Python package TaxoNERD (Le Guillarme & Thuiller 2022). Briefly, TaxoNERD is a deep learning-based named entity recognition tool pretrained on biomedical corpus using transfer learning. TaxoNERD extracts biological entities from text and associates to each one its NCBI Taxonomy ID. We then used the ETE Toolkit (Huerta-Cepas, Serra, and Bork, 2016) to extract the higher level taxonomic classifications from the NCBI database for each TaxID found by TaxoNERD and summarize the number of mentions for each species, genus, family, and order.

To identify sampling gaps in focal organisms, we compared the observed number of papers focused on each order of land plants to the number that we would expect if research attention had been evenly distributed across focal organisms relative to the species richness of the order. We identified orders that were statistically over- and under-represented with Fisher's Exact Tests. Summary stats, regional patterns, and changes over time were computed using Python packages Pandas and Numpy and visualized in Seaborn and Matplotlib.

DATA AVAILABILITY

Data associated with this study and a description of data acquisition and curation are deposited in Dryad at https://doi.org/10.5061/dryad.pg4f4qrtb.

AUTHOR CONTRIBUTIONS

RAM, EJA, SP, ARC, CCR, SMT, JMF, DHC and RV conceived of the study. RAM, EJA, SP, ARC, SMT, contributed to data acquisition and curation. RAM, EJA, and SP conducted data analyses. RAM, EJA, SP, ARC, CCR, SMT, JMF, DHC and RV contributed to data interpretation and conceptual framing of the manuscript. RAM, EJA, and SP drew the figures. RAM wrote the manuscript. All authors edited and reviewed the manuscript.

ACKNOWLEDGEMENTS

This work was supported by NSF Postdoctoral Research Fellowship in Biology (grant no. PRFB-1906094) to RAM and funding from the Michigan State University Department of Biochemistry and Molecular Biology to SP. Other members of our working group were supported by the USDA National Institute of Food and Agriculture, Michigan State University AgBioResearch, the National Research Foundation of South Africa, and Universidad Nacional Autónoma de México.

REFERENCES

Adamo, M., M. Chialva, J. Calevo, F. Bertoni, K. Dixon, and S. Mammola. 2021. Plant scientists' research attention is skewed towards colourful, conspicuous and broadly distributed flowers. *Nature plants*.

Adas, M. 2008. Colonialism and Science. *In* H. Selin [ed.], Encyclopaedia of the History of Science, Technology, and Medicine in Non-Western Cultures, 604–609. Springer Netherlands, Dordrecht.

Amarante, V., R. Burger, G. Chelwa, J. Cockburn, A. Kassouf, A. McKay, and J. Zurbrigg. 2021. Underrepresentation of developing country researchers in development research. *Applied economics letters*: 1–6.

Anon. World Economic Situation and Prospects 2020. Website https://www.un.org/development/desa/dpad/publication/world-economic-situation-and-prospects-2020/ [accessed 9 September 2022].

Ardila-Mantilla, F. 2016. Todos Cuentan: Cultivating Diversity in Combinatorics. *Notices of the American Mathematical Society. American Mathematical Society* 63: 1164–1170.

Baber, Z. 2016. The Plants of Empire: Botanic Gardens, Colonial Power and Botanical Knowledge. *Journal of contemporary Asia* 46: 659–679.

Bettencourt, L. M. A., J. Lobo, D. Helbing, C. Kühnert, and G. B. West. 2007a. Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the National Academy of Sciences of the United States of America* 104: 7301–7306.

Bettencourt, L. M. A., J. Lobo, and D. Strumsky. 2007b. Invention in the city: Increasing returns to patenting as a scaling function of metropolitan size. *Research policy* 36: 107–120.

Bettencourt, L. M. A., J. Lobo, D. Strumsky, and G. B. West. 2010. Urban scaling and its deviations: revealing the structure of wealth, innovation and crime across cities. *PloS one* 5: e13541.

Bonham, K. S., and M. I. Stefan. 2017. Women are underrepresented in computational biology: An analysis of the scholarly literature in biology, computer science and computational biology. *PLoS computational biology* 13: e1005134.

Breda, T., E. Jouini, C. Napp, and G. Thebault. 2020. Gender stereotypes can explain the gender-equality paradox. *Proceedings of the National Academy of Sciences of the United*

States of America 117: 31063–31069.

Carter, A. L., and A. Alexander. 2020. Soul food: [re]framing the African-American farming crisis using the culture-centered approach. *Frontiers in communication* 5.

Castro-Peraza, M. E., J. M. García-Acosta, N. Delgado, A. M. Perdomo-Hernández, M. I. Sosa-Alvarez, R. Llabrés-Solé, and N. D. Lorenzo-Rocha. 2019. Gender Identity: The Human Right of Depathologization. *International journal of environmental research and public health* 16.

Crenshaw, K. 1989. Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics. *The University of Chicago legal forum* 1989: 8.

Dirzo, R., G. Ceballos, and P. R. Ehrlich. 2022. Circling the drain: the extinction crisis and the future of humanity. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 377: 20210378.

Dwyer, W., C. N. Ibe, and S. Y. Rhee. 2022. Renaming Indigenous crops and addressing colonial bias in scientific language. *Trends in plant science*.

Ebenezer, T. E., A. W. T. Muigai, S. Nouala, B. Badaoui, M. Blaxter, A. G. Buddie, E. D. Jarvis, et al. 2022. Africa: sequence 100,000 species to safeguard biodiversity. *Nature* 603: 388–392.

Fox, C. W., C. S. Burns, A. D. Muncy, and J. A. Meyer. 2016. Gender differences in patterns of authorship do not affect peer review outcomes at an ecology journal. *Functional ecology* 30: 126–139.

Frances, D. N., C. R. Fitzpatrick, J. Koprivnikar, and S. J. McCauley. 2020. Effects of inferred gender on patterns of co-authorship in ecology and evolutionary biology publications. *Bulletin of the Ecological Society of America* 101.

Gilligan, Carol, Snider, and Naomi. 2018. Why does patriarchy persist?

Golash-Boza, T., M. D. Duenas, and C. Xiong. 2019. White Supremacy, Patriarchy, and Global Capitalism in Migration Studies. *The American behavioral scientist* 63: 1741–1759.

Haghighat-Sordellini, E. 2010. Patriarchy, Modernization, and the Global Economy. *In* E. Haghighat-Sordellini [ed.], Women in the Middle East and North Africa: Change and Continuity, 33–49. Palgrave Macmillan US, New York.

Hendre, P. S., S. Muthemba, R. Kariba, A. Muchugi, Y. Fu, Y. Chang, B. Song, et al. 2019. African Orphan Crops Consortium (AOCC): status of developing genomic resources for African orphan crops. *Planta* 250: 989–1003.

Holman, L., D. Stuart-Fox, and C. E. Hauser. 2018. The gender gap in science: How long until women are equally represented? *PLoS biology* 16: e2004956.

Ibe, C. N. 2022. Democratizing plant genomics to accelerate global food production. *Nature genetics* 54: 911–913.

Kocabıçak, E. 2022. The Political Economy of Patriarchy in the Global South. Taylor & Francis.

Lagakos, D. 2020. Does Internal Migration Offer Opportunities? *The journal of economic perspectives: a journal of the American Economic Association* 34: 174–192.

Larivière, V., C. Ni, Y. Gingras, B. Cronin, and C. R. Sugimoto. 2013. Bibliometrics: global gender disparities in science. *Nature* 504: 211–213.

Lerman, K., Y. Yu, F. Morstatter, and J. Pujara. 2022. Gendered citation patterns among the scientific elite. *Proceedings of the National Academy of Sciences of the United States of America* 119: e2206070119.

Li, F.-W. 2021. Decolonizing botanical genomics. *Nature plants* 7: 1542–1543.

Maas, B., R. J. Pakeman, L. Godet, L. Smith, V. Devictor, and R. Primack. 2021. Women and Global South strikingly underrepresented among top‐publishing ecologists. *Conservation letters* 14.

Madzima, T. F., and G. C. MacIntosh. 2021. Equity, diversity, and inclusion efforts in professional societies: intention versus reaction. *The Plant cell* 33: 3189–3193.

Maril, R. K. & Gill, A.M. 2018. Discrimination Based on Perceived Characteristics. Human Rights Campaign Foundation, Washington, DC.

Marks, R. A., S. Hotaling, P. B. Frandsen, and R. VanBuren. 2021. Representation and participation across 20 years of plant genome sequencing. *Nature plants*.

Montgomery, B. L. 2020. Academic leadership: gatekeeping or groundskeeping? J Values-Based Leadersh 13: 135--151.

Montgomery, B. L., and J. A. Whittaker. 2022. The roots of change: Cultivating equity and change across generations from healthy roots. *The Plant cell* 34: 2588–2593.

Munstermann, M. J., N. A. Heim, D. J. McCauley, J. L. Payne, N. S. Upham, S. C. Wang, and M. L. Knope. 2022. A global ecological signal of extinction risk in terrestrial vertebrates. *Conservation biology: the journal of the Society for Conservation Biology* 36: e13852.

Santamaría, L., and H. Mihaljević. 2018. Comparison and benchmark of name-to-gender inference services. *PeerJ. Computer science* 4: e156.

Smith, R. J. 2009. Use and misuse of the reduced major axis for line-fitting. *American journal of physical anthropology* 140: 476–486.

Stefanoudis, P. V., W. Y. Licuanan, T. H. Morrison, S. Talma, J. Veitayaki, and L. C. Woodall. 2021. Turning the tide of parachute science. *Current biology: CB* 31: R184–R185.

Stoet, G., and D. C. Geary. 2018. The Gender-Equality Paradox in Science, Technology, Engineering, and Mathematics Education. *Psychological science* 29: 581–593.

Szövényi, P., A. Gunadi, and F.-W. Li. 2021. Charting the genomic landscape of seed-free plants. *Nature Plants*: 1–12.

Topaz, C. M., J. Higdon, A. Epps-Darling, E. Siau, H. Kerkhoff, S. Mendiratta, and E. Young. 2022. Race- and gender-based under-representation of creative contributors: art, fashion, film, and music. *Humanities & social sciences communications* 9: 221.

Trimble, M. J., and R. J. van Aarde. 2012. Geographical and taxonomic biases in research on biodiversity in human-modified landscapes. *Ecosphere* 3: art119.

Trisos, C. H., J. Auerbach, and M. Katti. 2021. Decoloniality and anti-oppressive practices for a more ethical ecology. *Nature Ecology & Evolution*: 1–8.

Uchendu, E., and N. Edeagu. 2021. Negotiating Patriarchy and Gender in Africa: Discourses, Practices, and Policies. Rowman & Littlefield.

Uylings, H. B., C. G. van Eden, and M. A. Hofman. 1986. Morphometry of size/volume variables and comparison of their bivariate relations in the nervous system under different conditions. *Journal of neuroscience methods* 18: 19–37.

Vallée, G. C., D. S. Muñoz, and D. Sankoff. 2016. Economic importance, taxonomic representation and scientific priority as drivers of genome sequencing projects. *BMC genomics* 17: 782.

Wapman, K. H., S. Zhang, A. Clauset, and D. B. Larremore. 2022. Quantifying hierarchy and dynamics in US faculty hiring and retention. *Nature*.

West, J. D., J. Jacquet, M. M. King, S. J. Correll, and C. T. Bergstrom. 2013. The role of gender in scholarly authorship. *PloS one* 8: e66212.

Witteman, H. O., M. Hendricks, S. Straus, and C. Tannenbaum. 2019. Are gender gaps due to evaluations of the applicant or the science? A natural experiment at a national funding agency. *The Lancet* 393: 531–540.