Environmental Science and Policy 126 (2021) 79–89

Contents lists available at ScienceDirect

Environmental Science and Policy

journal homepage: www.elsevier.com/locate/envsci

Selective world-building: Collaboration and regional specificities in the marine biodiversity field

Petro Tolochko [a], Alice B.M. Vadrot [b,*]

[a] University of Vienna, Department of Communication, Austria
[b] University of Vienna, Department of Political Science, Austria

ARTICLE INFO

Keywords:
Marine biodiversity
Scientific collaboration
Ocean science
Bibliometric analysis

ABSTRACT

Over the past two decades, the scientific study of marine biodiversity developed into one of the most dynamic research fields within environmental research in general and ocean science in particular. Marine biodiversity research spans a broad range of spatial and temporal scales, scientific disciplines, and infrastructures for assessing patterns of change in marine biodiversity with different topics and regions contributing to the abstract construction of the field. Yet, beneath the surface of these abstract constructions and attributions such as “big science” to assess the international, interdisciplinary and data-driven nature marine biodiversity research, patterns of scientific collaboration emerge that shape the contributions of different world regions to the field. Furthermore, legal, political and territorial orders shape scientific collaboration by determining access to the marine environment and the study thereof. To understand these dynamics, this study analyses scientific publications on marine biodiversity using topic modeling methodology. Our approach provides the opportunity not only to describe large clusters within the field of marine biodiversity, but also to show how these topical clusters differ in time and across regions. By looking at the temporal and regional level, we wish to contribute to a deeper understanding of collaboration and regional specificity in the emergence of new research fields and – what we call – “selective world-building” across time and space.

1. Introduction

Marine biodiversity is one of the most dynamic research fields within environmental research in general and ocean science in particular spanning a broad range of spatial and temporal scales for assessing patterns of change in marine biodiversity across the globe. While the term “biodiversity” first appeared in the mid-1980s (Wilson, 1985), it only gained serious traction in the 1990s and has continued to increase in its popularity even today (Vadrot, 2020). Marine biodiversity is often thought of as a “big science” (Price, 1986), characterized by big investments, big research teams, big goals, and big international research collaborations and longitudinal monitoring programs (Vermeulen, 2013). This makes marine biodiversity a data-intense science, where different disciplines and data collection techniques are drawn together to produce coherent representations of the field (Bowker, 2000). Furthermore, and through a combination of different factors - such as increased demands for usable knowledge to protect marine ecosystems by policymakers and struggle over international regulations of ocean science- marine biodiversity research firmly established itself (as a field) in a political dimension.

Some of the factors that propelled marine biodiversity into politics are the aforementioned “big science” nature of the field – presupposing access to large and expansive research infrastructure and ocean exploration and data collection sites. These dynamics are closely tied to the legal, political, and territorial orders determining access to the marine environment and the study thereof. Different legal orders and principles apply to different maritime zones with implications for the planning and design of research expeditions, data collection and management practices, and the ownership of related research outcomes and materials. Furthermore, unequal access to research infrastructure and equipment perpetuates deeply rooted inequalities between the Global North and the Global South in terms of exploration and exploitation of marine resources, most notably in areas beyond national jurisdiction and the deep sea (Tolochko and Vadrot, 2021; Rogers et al., 2021).

International scientific collaboration and a division of work between scientists from different regions of the world and scientific disciplines, including “physical, biological, chemical, geological, hydrographic, health and social sciences, as well as engineering, the humanities and

---

* Corresponding author.
E-mail address: alice.vadrot@univie.ac.at (A.B.M. Vadrot).

https://doi.org/10.1016/j.envsci.2021.09.003
Received 3 February 2021; Received in revised form 16 July 2021; Accepted 3 September 2021

multidisciplinary research on the relationship between humans and the ocean" are central pillars of ocean science (Valdes, 2017). Collaborative research between regions and disciplines is central for advancing baseline information, systematic monitoring, and basic research of marine biodiversity patterns and processes. Yet, previous studies have also shown that scientific collaboration may have unintended effects, such as the continuation of colonially shaped science systems in tropical marine science (Partelow et al., 2020) or the rather low level of regional collaboration on marine biodiversity related issues between Asian, Latin American, and African countries (Tolochko and Vadrot, 2021). There is, however, a gap in understanding if – and if yes, how – scientific collaboration has shaped marine biodiversity research and the topical, temporal, and spatial specificities of the field, especially since the beginning of the 1990s when the first scientific papers published on "marine biodiversity" started their journey into the constitution of one of the most dynamic fields of environmental research.

One entry point to unpack the field of marine biodiversity is to focus on scientific publications and text. Through the joint production of scientific text scientists collectively engage in the abstract construction of scientific fields, or – as (Callon et al., 1986) have framed it – in "building a structured world" through text. For Callon, activities that are central to scientific collaboration such as the sharing of highly specialized equipment to carry out research, the composition and organization of research teams, the acquisition of funding, and strategies for securing access to research objects are central to the specific dynamics of a research field become observable objects in the guise of scientific publications (Callon et al., 1986). Yet, the building of structured worlds does not correspond to objective representations of a field; instead such processes create and manifest selective worlds.

The current paper aims to be one of the first attempts at a holistic analysis of the marine biodiversity field by drawing on the idea of "world-building" as developed by (Callon et al., 1986), which we use as a conceptual framework to study the dynamics of scientific fields through a bibliometric analysis of scientific texts. In this paper, we attempt to investigate what exactly constitutes the field of marine biodiversity in terms of content, to understand what regions play what role in the production of scientific knowledge within the field, and finally to assess whether international collaboration has an effect on the thematic focus of research.

Our study is relevant in many ways, most notably because of the increased interest in marine biodiversity as a research object, a resource to be protected, and a commodity to be indexed and used. First, the UN-Decade of Ocean Science estimates that 95% of marine species still need to be classified, calling for increased international efforts to close data gaps and increase our knowledge base on marine biodiversity (Sierra-Correa et al., 2020). Second, an economic interest in marine biodiversity, especially in marine genetic resources (MGRs), has been growing, leading to an increase in patents on MGRs and new discoveries (Blasiak et al., 2018). Third, the increasing privatization of the sea is aggravated by the lack of political regulations in the High Seas, covering two-thirds of the world ocean. And finally, marine biodiversity protection and sustainable use, especially in areas beyond national jurisdiction, are high on the political agenda, with ongoing negotiations for a new international treaty under the United Nation Convention on the Law of the Sea (UNCLOS) and recent assessment reports, indicating increased pressures on the marine environment (Brondizio et al., 2019; United Nations, 2021).

2. Viewing the'marine biodiversity field' through the lens of scientific collaboration and "selective world building"

The term "biodiversity" – including both biological diversity in the ocean and on land- was first coined in the mid-1980s (Wilson, 1985). As a "unifying imperative", it united scientists and practitioners holding widely different views and sometimes competing values regarding the conservation of the natural world (Takacs, 1996). Despite significant increase in interest in biodiversity as a field at the intersection between science and policy (Vadrot, 2014), almost no attention was placed to the field of marine biodiversity and how it has been shaped by dynamics within the research field, such as international scientific collaboration.

2.1. *Scientific collaboration within marine biodiversity research*

Literature from the natural sciences shows that the field of "marine biodiversity" is simultaneously considered a field in and of itself (e.g., Cochrane et al., 2016), a part of ocean science (e.g., Pendleton et al., 2020), and a part of a larger biodiversity science (Wilson, 1988). Marine biodiversity research could be described as a "big science" (Price, 1986), characterized by interests not only in regional but global phenomena, use of expensive large scale physical infrastructure, scientific collaboration, interdisciplinary nature, and large scale international research programs and observing systems (e.g., Global Ocean Observing System, Ocean Biogeographic Information System, Census of Marine Life etc.).

A significant amount of marine biodiversity research would not only be extraordinarily difficult to achieve by a single research lab but rather impossible without international and interdisciplinary collaboration. The international character of marine biodiversity is also dictated by a very complex infrastructure that has to be established in order to effectively coordinate such large scale projects, thus creating an international network that is tightly connected out of necessity (Bowker, 2000). These large-scale international projects not only tightened international collaboration but also underlined the different specializations of collaborating regions, furthering the research on marine biodiversity through a synergy of expertise (Vermeulen, 2013). In marine biology, for instance, the use of genetic technologies to classify marine species did not fully replace classical taxonomy, because of technical difficulties to apply the barcode system to specific species of fish (Vermeulen, 2013).

While such globalized views may be useful to capture important dynamics within scientific fields, they go at the expense of a more granular perspective and how such dynamics are shaped by developments in different regions of the world and collaboration between them. (Partelow et al., 2020), for instance, demonstrate that tropical marine science is significantly shaped by a division of labor between the global North having the capacities to conduct ocean science and the global South facilitating access to its data objects rather than being a research subject on its own. By investigating the structure of international collaboration the authors detect historical path dependencies hindering researchers and knowledge holders in the tropics to develop their ontologies for studying their marine environments (Partelow et al., 2020). Furthermore, past studies show that regional biodiversity research networks exhibit a different level of development, which the majority of regional research networks focusing on North American and European collaboration (Tolochko and Vadrot, 2021). Other bibliometric endeavors with a bigger focus on regional topical output – which could be thought of as regional "specializations" – look at the biodiversity field as a whole (Tydecks et al., 2018), but, as the same study notes, the vast majority of research in their sample consists of research on terrestrial biodiversity systems (roughly 85%), leaving little room for the actual marine biodiversity field. The lack of systematic investigation of the emerging marine biodiversity field and the temporal and spatial specificities of global and regional dynamics of the field is the starting point of this paper.

3. Science is politics by other means: Selective World building

One entry point for studying the emergence of the marine biodiversity field and how it has been shaped by temporal and spatial specificities is to survey its dynamics based on scientific publications and text. According to Callon, texts play a particular role in science and in the attempts of scientists in "building a structured world" (Callon et al., 1986, p. 9).

Activities such as the acquisition of funding, the management of laboratories, scientific staff, and equipment, scientific collaboration, and the division of work are all built into scientific publications. The specificities of "big science" involve the organization of big equipment and many scientists that tie different bodies of expertise into a common research project. This leads to multi-authored publications, with long lists of funding bodies that supported different researchers, research tasks, and equipment. Making visible the relation to others is inherent to scientific publications that contain references to literature, other scientists, and institutions (Callon et al., 1986).

One key assumption that underpins this view is that science is politics by other means (Callon et al., 1986; Latour et al., 1983). The production of science and scientific texts are social practices conditioned by a specific political context (Pielke, 2004; Sarewitz, 2004). Even in cases where scientific discourses may seem technical, such as in the case of the sixth mass extinction or invasive species, they are implicitly political (Vadrot, 2020; Warren, 2007; Litfin, 1994). Governments can influence the development of scientific disciplines and research directions through resource allocation, which – in the case of "big science" projects – culminate into large-scale investments into science and data infrastructures. (Skolnikoff, 2001), for instance, notes that "publicly funded scientific research is still national in nature, with budgets, goals, and choices determined almost wholly in the context of national decision-making processes", while (Hallonsten, 2014) notes that large-scale research programs (in the case of the European Union) require a shrewd political decision-making from collaborating scientists. Science is politics by other means because it continues to systematically sideline weaker actors with fewer capacities to contribute to large-scale research projects.

The idea of selective world-building, as described by Callon et al. allows us to conceptualize how scientific text contributes to both the creation of consistency despite diversity within scientific fields and, the production of selectivity through regional research choices and capacities. Publications tie scientific collaboration, new technologies, investments, but also a division of labor (those who collect the data, those who analyze the data, those that give access to data, and those who pay for research) together, but this does not necessarily imply coherence between regions and may lead to selective representations of the field.

In the current paper, we take a scientific publication as a basic building block of international scientific collaboration and aim to investigate how scientific collaboration within the marine biodiversity field is shaped by regional and topical specificities.

4. Mapping the dynamics of a research field with bibliometric analysis

Bibliometric analysis encompasses a range of techniques that are used in order to study and understand the output or impact of published scientific work on all levels, ranging from an individual researcher, an institution, or the whole scientific field, and is considered to e one of the main ways to understand and describe the activity of these actors (Zitt, 2005). According to (Noyons et al., 1999a) there are two main paradigms in bibliometric research – performance analysis and science mapping. The former deals with the evaluation of actors (single researchers, institutions, countries, etc.), while the latter deals with analyzing the structure of a scientific field. One of the key problems in the science mapping paradigm is to understand what actually constitutes a "research field" and how to delimit it (Leydesdorff and Rafols, 2009; Noyons et al., 1999b; Small, 2003; Upham and Small, 2010).

Two quantitative bibliometric techniques are being used to create 'science mapping' (i.e., a structural representation of a research field) – citation analysis (Calero-Medina and Noyons, 2008; Small, 1977), in which researchers aim to understand the scientific field through documents that are frequently cited together, and word co-occurrence analysis, in which the information about the research field is being extracted from the association of words that being used in different documents (Callon et al., 1983; Cobo et al., 2011). The terms "science mapping" and "field delimitation" are often synonymous and are used to "determine its (cognitive) structure, its evolution, and main actors within" (Noyons et al., 1999a). The subdomains or subfields are thus 'cognitively related core keywords in the studied field'. However, the majority of carried out bibliometric research tends to focus only on a subset of a given research field, or on the measurement of the performance of actors in the field, while the conceptual analysis (both quantitative and qualitative) of the field as a whole and its overarching structure remains rather limited (Cobo et al., 2011). Similar approaches have been taken to study various scientific fields ranging from management science (Ramos-Rodríguez and Ruíz-Navarro, 2004) to software engineering (Coulter et al., 1998) among others.

The current paper takes the approach of word co-occurrence analysis to define the research field of "marine biodiversity". However, rather than relying on the simple word co-occurrences to delimit subfields within the field of Marine Biodiversity, we approach this problem using state-of-the-art text-as-data methodology such as structural topic modeling (Roberts et al., 2014a). There are several good theoretical reasons to approach an analysis of the structure of a scientific field with a topic modeling methodology. According to (Börner et al., 2003; Callon et al., 1991; Cobo et al., 2011), the co-occurrence of keywords from scientific publications shows the similarity of between these publications, and if several keywords are connected through and can show a group of documents related to the specific scientific field. The co-occurrence can further be fine-grained to discern specific subfields within an overarching field. The topic modeling approach encompasses the idea that fields and subfields are formed via a connected cluster of words, but rather than simply relying on single keywords from a particular study, more textual information can be used to classify the subfields. From word co-occurrences within a single document (text) and the whole corpus, the model infers different "topics" – thematic clusters of specific word co-occurrences and then tries to infer the distribution of these topics over the text. This provides significantly more information to infer the 'topical clusters' than using keywords and allows for additional flexibility because a single document may incorporate several thematic topics.

Theoretically, this approach to analyzing subdomains within a given field is quite faithful to the one described above by (Noyons et al., 1999b), with the exception that instead of defining subdomains as a densely connected cluster of keywords, we define them as separate topics from the topic model (or, as will be evident from the Method section, as highly correlated clusters of topics). Additionally, topic modeling approaches have large advantages over word co-occurrence analysis especially in large corpora (Leydesdorff and Nerghes, 2017), which is the case of the current study.

Importantly, structural topic models (Roberts et al., 2014a,b), serve not only as a simple clustering algorithm for textual data but also allow including different covariates in topical estimation. This provides an opportunity to analyze not only subfields associated with the field of Marine Biodiversity but also their distribution across time and regions. Effectively providing the ability to see the evolution of Marine Biodiversity (and its subfields) for each individual region.

5. Methodological approach

In the interest of space, the section below is only a short description of the methodology used in the current article. For more detail about the methods employed and the analyses procedure, please refer to the Supplementary Material.

5.1. Data and sample

The data used in the study were collected from the Web of Science digital abstract repository. The articles were related to the "marine biodiversity" field and were indexed using the following search string:

*TOPIC: (”marine *diversity”) OR (biodivers* (marine OR sea OR ocean)) OR (deep-sea biodivers*) OR (”marine protected area*” biodivers*) OR (sea biodivers*) OR (deep-ocean* biodivers*) OR (”genetic diversity” (marine OR sea OR ocean)) OR (”species diversity” (marine OR sea OR ocean)) OR (”taxonomic divers*” (marine OR sea OR ocean)) OR (”ecosystem divers*” (marine OR sea OR ocean)) OR (”functional divers*” AND (marine OR sea OR ocean)).*

In total, the dataset comprised of 24,219 scientific articles published in the window from 1990 to 2018, from 153 countries (based on the first-author classification), with 21,919 (90.5%) being multi-authored and 9700 (40%) being international (i.e., being written by authors from two or more different countries). The abstracts came from 2711 scientific journals and conference proceedings. The sample only included peer-reviewed publications, excluding gray literature. While the concept and term "marine biodiversity" appears in academic abstracts before 1990, this is the first year in our data where this term appears systematically in consecutive years; a decision was taken to start the analyses from the year 1990.

5.2. Topic modeling

Several topic models with were estimated with K (number of topics) ranging from 60 to 90. After comparing held-out likelihoods and semantic coherence (Wallach et al., 2009) of the models, and, more importantly, a qualitative inspection of each model, a K = 70 model was chosen as optimally suitable for the current dataset/research design. Furthermore, the 70 separate topics were clustered into 17 large clusters to facilitate the interpretation of results. These clusters represent a collection of topics from the topic modeling output, which, in turn, are a collection of strongly associated words (grouped together based on a hierarchical cluster analysis). Online supporting information contains a table in which we show which words relate to which topics, and which topics to which clusters. Thus, a cluster, as used in this paper, is a general thematic focus into which scientific publications fall.

5.3. Pairwise collaboration

In order to later be able to analyze how topic proportions vary across collaborative and non-collaborative articles from different regions, a series of variables were created to encode information on pairwise collaboration. In order to read the specifics of how pairwise collaboration was determined, please refer to the Supplementary Material.

5.4. Analysis

The topic model was estimated with the year of publication and region of the first author as covariates, i.e., the topic proportions are allowed to vary across time as well as geographically. After the initial model was estimated and selected (see the section above), and topics and clusters labeled, the model was used to re-estimate the proportion of clusters across collaboration covariates, and the structural topic model estimation function was modified to allow for “cluster prevalence”, instead of topic prevalence to vary across the covariates.

5.5. Collaboration regression

In order to investigate the general patterns of collaboration across different regions, we have modeled the number of collaboration articles between different regions as a function of regions’ expertize in a particular cluster. The region’s *Expertise* on a specific cluster was used as an independent variable, and the *Collaboration Rate* on that cluster as the dependent variable. We have additionally included several control variables (some were included as random effects in a multilevel model). *Log-transformed total number of publications* from Region A and Region B were included as covariates into the model to control for both regions’ productivity. Random effects were clustered over *Thematic Clusters* and *Undirected Region Pairs* (i.e., Region A/Region B and Region B/Region A was treated as the same pair) to control for pairwise regional behavior. Please refer to Supplementary Material for further information.

6. Specificities and dynamics of the marine biodiversity field

In the process of analyzing the output of the topic model as well as subsequent clustering efforts, we have identified 17 separate thematic clusters that together constitute the field of marine biodiversity according to the bibliometric data obtained from Web of Science Core Collection from 1990 until 2018. The 17 clusters, as well as their respective proportions across the whole corpus, are provided in Table 1.

Some clusters are straightforward to interpret - e.g., *Methods* (mapping, monitoring, prediction, classification, modeling, and indicators etc.), *Fishery*, (abundance of biomass, fisheries, catch, fish stocks management etc.), or *Tropical diversity/archipelago* (biogeography, species diversity, origin, patterns in tropical environments and islands), *Land-sea relations* and *Sediment/deep sea*, which are characterized by spatial and geographical aspects. Other clusters and cluster groups are less obvious, such as the clusters related to genetics. *Genetic Diversity* is a cluster devoted to genetic variation in populations, species, or organisms, while *Gene Sequencing* is mostly related to DNA sequencing and its use for product development based on marine genetic resources. The distinction between *Genetic Taxonomy* and *Classical Taxonomy* is characterized by methodology - the use of genomic typing or observable morphological characteristics. *Environmental Management* is, as the name implies, concerned with, monitoring and managing environmental change, as well as biodiversity conservation, restoration, and sustainable use, including marine protected areas, ecosystem services, policy, governance and review.

We provide additional information on the selected clusters below, however in the interest of space, the relationship between all clusters and the corresponding academic abstracts that they represent is provided in Online Supporting Information.

The two most prominent clusters are gene-related (i.e., *Genetic Diversity* and *Gene Sequencing*), the third most popular cluster in the complete corpus is the *Environmental Management* cluster, although this is a highly region-specific position. The methodological cluster is also quite high in the ranks, indicating a strong methodological focus of the marine biodiversity field.

Interestingly, there are two separate *Taxonomy* clusters outlined in the corpus – *Genetic Taxonomy* and *Classical Taxonomy*. While *Genetic Taxonomy* is a fairly popular cluster in the field of marine biodiversity, *Classical Taxonomy* is occupying only a marginal spot in the cluster distribution. It is important to note, that apart from the first two genetic

Table 1
Identified clusters with absolute and relative numbers.

	Cluster Name	Absolute	Prop. (%)
1	Genetic diversity	3944	16.28
2	Gene, product, sequence	2882	11.90
3	Environmental management and protection	1863	7.69
4	Methods	1762	7.28
5	Data, mass extinction	1462	6.04
6	Taxonomy (genetic)	1419	5.86
7	Ecosystem health and water quality, invasive	1402	5.79
8	Species and functional diversity, community	1387	5.73
9	Tropical diversity, archipelago	1201	4.96
10	Coral reef, sponge	1119	4.62
11	Algae & food web	991	4.09
12	Sediment, deep sea	969	4.00
13	Land sea relations	943	3.89
14	Taxonomy (classical)	782	3.23
15	Plankton, inland water, offshore	735	3.03
16	Climate change	686	2.83
17	Fish, fishery	672	2.77

Note: Absolute amount of articles related to a cluster, as well as proportional to the whole corpus.

clusters, which are significantly more frequent in the dataset, the frequency distribution of further clusters decays rather slowly, with the vast majority of clusters falling roughly between 7% and 4% of the whole dataset. Nevertheless, as will become evident from the next sections, these clusters are distributed differently both in the spatial and temporal dimensions.

6.1. Cluster distribution

We first present the general distribution of clusters in the whole dataset, as well as broken down by regions. Fig. 1 shows the relative frequency of clusters in the complete dataset as well as broken down by regions. As is evident from Fig. 1, the distribution of clusters is very non-uniform, with the two top clusters (*Genetic Diversity* and *Gene Sequencing* being by far the most popular in the full dataset. The highest frequency cluster – *Genetic Diversity* remains the most popular for all regions, indicating that the topical interest in marine biodiversity science, at least on the most general level, remains consistent on the global scale.

We further show the temporal distribution of marine biodiversity clusters. Cluster prevalence over time was analyzed using a cubic spline model with 12 knots. In general, clusters stay relatively stable across the analyzed time frame (i.e., from 1990 until 2018). Nevertheless, one particular cluster – *Environmental Management and Protection* – shows a distinct rise in popularity in our dataset. Fig. 2 shows the top 5 clusters by popularity with cluster prevalence plotted against time. Cluster prevalence, unlike raw proportions seen in Fig. 1, refers to the proportion of clusters within texts. A certain cluster may always have the *highest* proportion, but these proportions change over time, indicating varying preferences of regions to focus (or shift focus from) certain clusters. As can be seen from the figure, the *Environmental Management* cluster develops an upward trend around years 2006–2008, while, around the same time, the *Data and Mass Extinction* cluster begin to slope down. Additionally, we can see that around the years 2010–2011, *Genetic Diversity* – the most frequent cluster in the dataset – also begins to develop a downward trend. These dynamics may indicate a shifting interest in marine biodiversity science from the genetic towards a more environmental management approach. Unfortunately, the sparsity of the data at the beginning of the analyzed time frame (from 1990 until roughly 2000) does not allow making meaningful comparisons between cluster popularity, as indicated by very wide credibility intervals.

6.2. Regional differences

We now shift the discussion from the general descriptives to region-specific differences. In general, as can be seen from Fig. 1, there are two constant clusters that are always on top regardless of the region – *Genetic Diversity* and *Gene Sequencing.* Further down, however, there is significant variability in cluster proportions across different regions. For example, Asia has a very strong presence in clusters related to genetics, which is exemplified by a significant amount of papers not only from the top 2 most frequent clusters (that deal with genetics) but also by the fact that *Genetic Taxonomy* is very prominent.

Nevertheless, marine biodiversity science is not as homogeneous as one might think. Fig. 3 shows relative cluster prevalence rather than raw count scores as in Fig. 1; unlike the raw count of the documents dedicated to a particular cluster, cluster prevalence shows the relative proportion of text dedicated to the topic, as well as showing the regional effect on this proportion. In essence, Fig. 3 shows the proportion of the text devoted to a cluster. *Genetic Diversity* and *Gene Sequencing* are the most prominent clusters, they are the most prominent across all regions, but we are also interested in how different clusters are represented in each region in relation to the whole corpus.

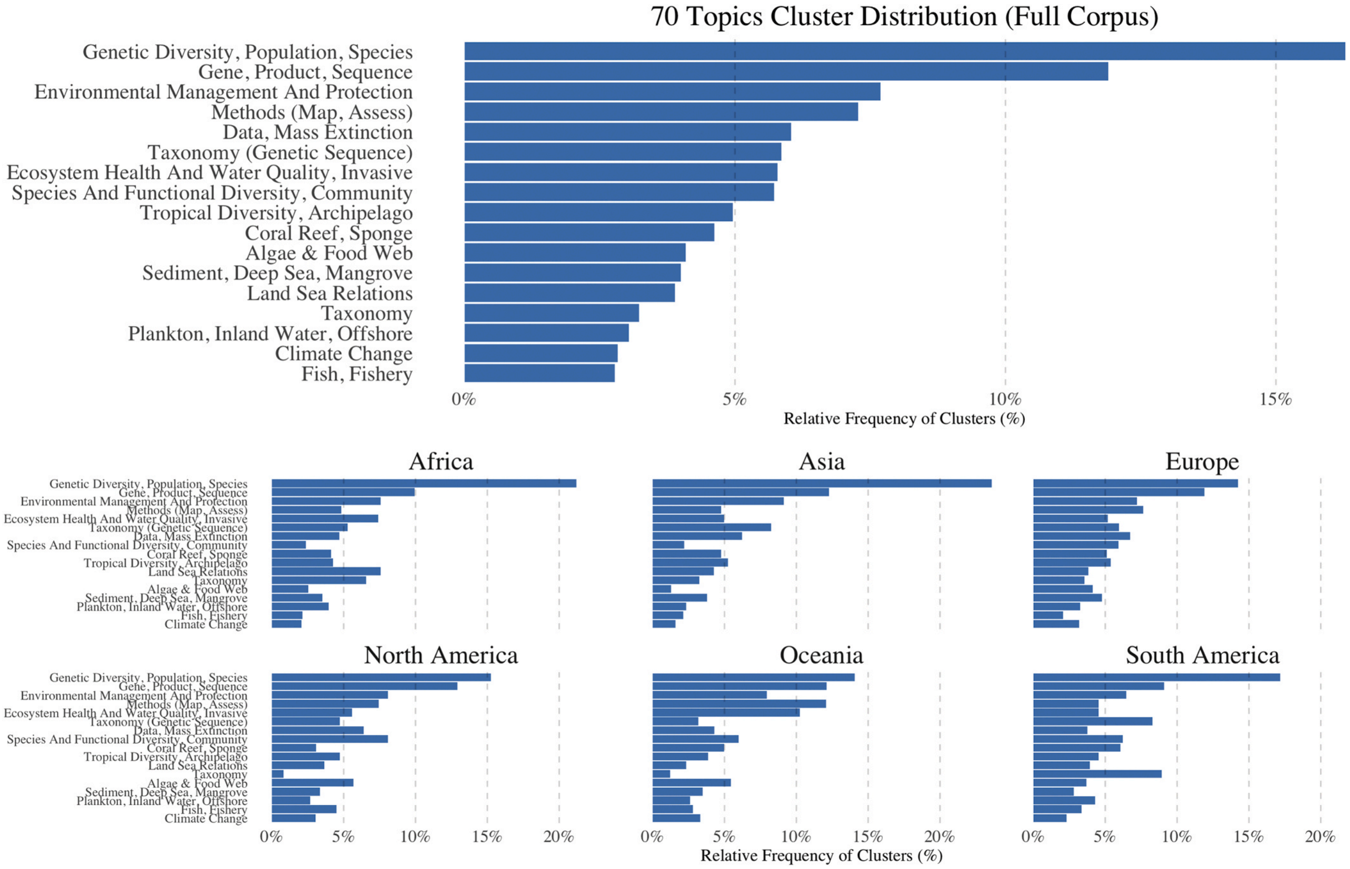

Fig. 1. Relative Frequency of Clusters. Top panel shows the relative frequency of clusters in the whole dataset. Bottom panel shows the relative frequency of clusters broken down by regions.

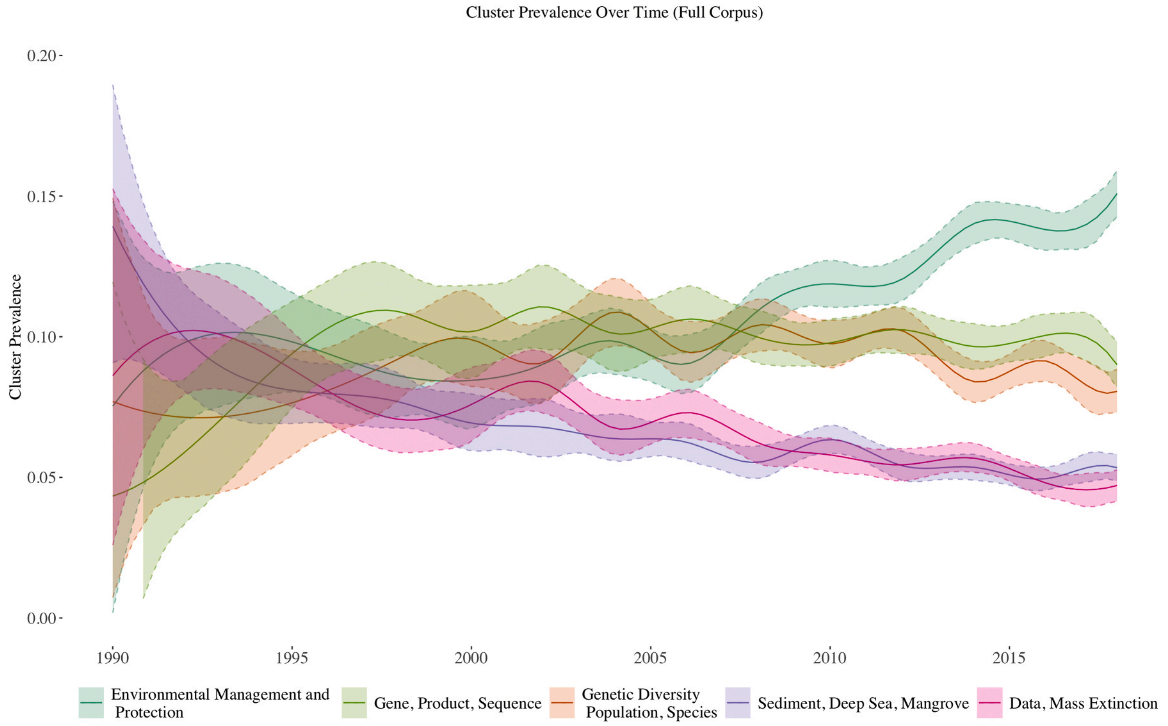

Fig. 2. Cluster Prevalence Over Time. Spline interpolation with cubic spline and 12 knots. Shaded regions show 90% Posterior Credibility Intervals. Note the width of Credibility Intervals at the beginning of the time frame, indicating very sparse data.

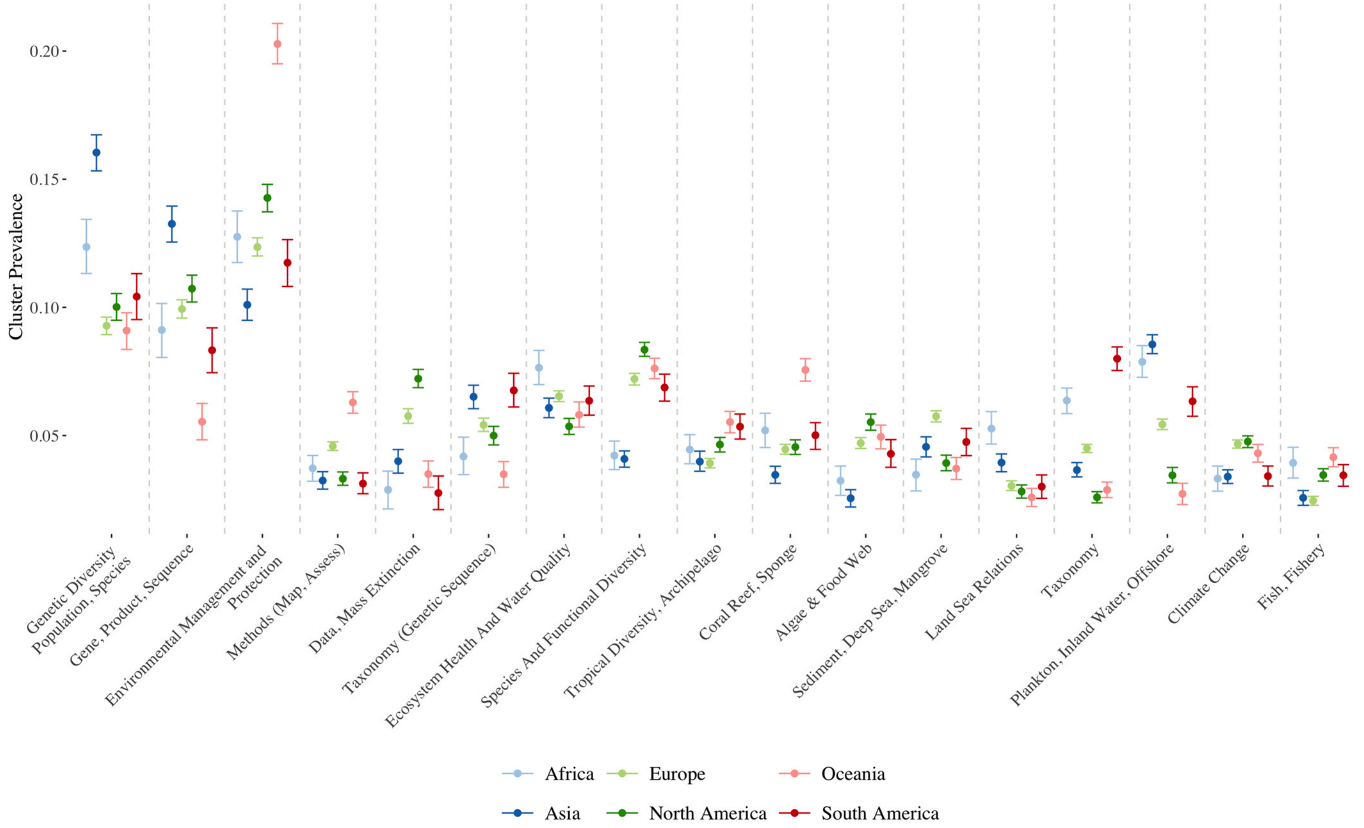

Fig. 3. Cluster Prevalence by Region. The x-axis shows clusters in order of relative frequency across the whole dataset. The y-axis shows the prevalence of specific regions, along with 90% Credibility Intervals.

Several regional patterns observed; Asia is much more likely to produce texts that are aligned with genetic sequencing (in one way or another) – *Genetic Diversity, Gene Sequencing* and *Genetic Taxonomy*, and is less prevalent (in comparison) in the *Environmental Management* cluster. Conversely, authors from Oceania tend to be less present in clusters related to genetics, however, they are significantly more present in the *Environmental Management* cluster compared to all other regions. Additionally, *Coral Reefs* and *Methodological* clusters are also significantly related to Oceania. On the other hand, South America and, to some extent, Africa, focuses on *Classical Taxonomy*, but are less productive in *Genetic Taxonomy*. Fig. 4 shows cluster prevalence plotted over time for each individual region. Since individual clusters have a significant amount of uncertainty associated with their distribution across time (this is especially evident for regions, where data is more sparse, such as South America and Africa), it is rather difficult to see the change in prevalence associated with time. Nevertheless, several observations could be immediately made from the plot. For the majority of regions, a clear rise in the popularity of the *Environmental Management* cluster is seen. This is especially evident in the regions of the global North (i.e., North America, Europe, and Oceania), where roughly after 2010 this becomes the most prevalent cluster (after 2005 in Oceania). However, the rise of this cluster in Asia is also observable, while it has never occupied the top position prevalence-wise, we can see a slight upward trend from about 2010 onward. The second most prevalent cluster, namely *Genetic Diversity* stays relatively constant from the beginning of our dataset in 1990 until 2018, except Asia. In Asia, this cluster was dominating the marine biodiversity field from around 2000, however, it began to see a constant downward trend around 2008, and by the end of 2018, its prevalence is on par with *Gene Sequencing* and *Environmental Management* clusters. Interestingly, the downward trend associated with *Genetic Diversity* started around the same time as the upward trend of *Environmental Management*.

6.3. Regional collaborations

To further understand the field of marine biodiversity, cluster prevalence in pair-wise cluster collaborations is calculated for all regions. Figures A1 through A6 show collaboration patterns of North America, Oceania, South America, Asia, and Africa, respectively (provided in the Appendix for the interest of space). The figures show which of the clusters are more prevalent in collaborative articles (red dots), less prevalent in collaborative articles (blue dots), or those that have no difference between collaborative and non-collaborative articles (gray dots). In order to have a clearer picture of what topics are, Fig. 5 shows stacked frequencies of clusters that are either more/less prevalent in collaboration articles or have no difference. The distribution of clusters from Fig. 5 shows that clusters that are more focused on specific regional themes (such as *Inland Waters, Food Web*, and *Taxonomy*) and economic clusters (such as *Genetics/Product* cluster) are being less collaborated on (i.e., their prevalence is higher in non-collaborative articles), while clusters that deal with global themes (such as *Environmental Management and Protection*, *Functional Biodiversity* and *Genetic Diversity*) are being more collaborated on.

As can be seen from Table 2, the estimate of ΔExpertise is negative and statistically significant, which indicates collaboration on a cluster is determined by the co-author's expertise on the topic, rather than the first author's expertise (i.e., when Region A and Region B are collaborating on Topic X, it is much more likely that this collaboration has occurred because of Region B's expertise). While the effect might seem small, it is important to remember that we are dealing with regional collaboration patterns; since the effect is presented as logged odds, on the probability scale this implies that $\exp(-0.05) = 0.95$. Thus, for every unit increase in Region B's expertise, there is a 5% chance increase in the probability that one additional collaboration paper with Region B is being published, which, on a regional level is not an insignificant effect.

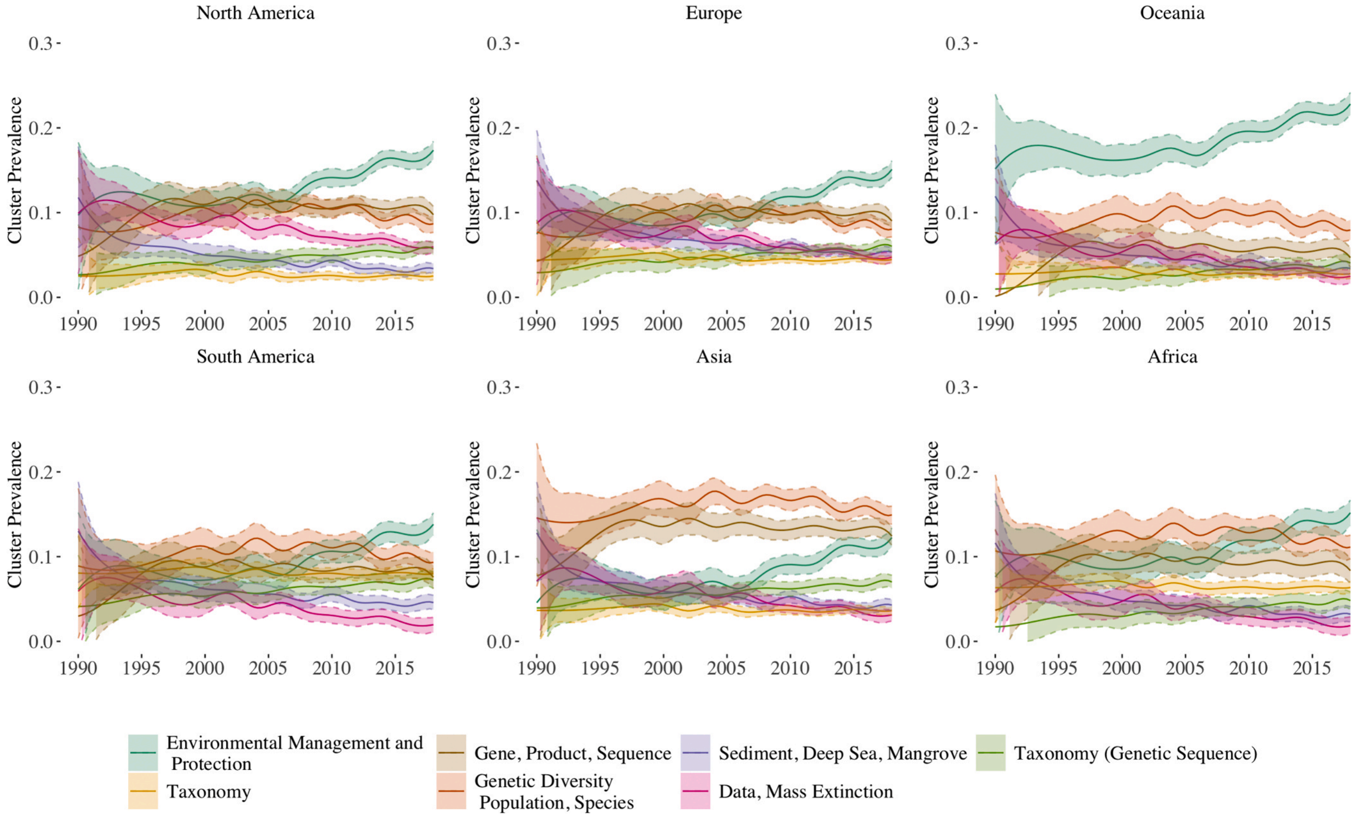

Fig. 4. Cluster Prevalence Over Time for individual regions. Spline interpolation with cubic spline and 12 knots. Shaded regions show 90% Posterior Credibility Intervals. Note the width of Credibility Intervals at the beginning of the time frame, indicating very sparse data.

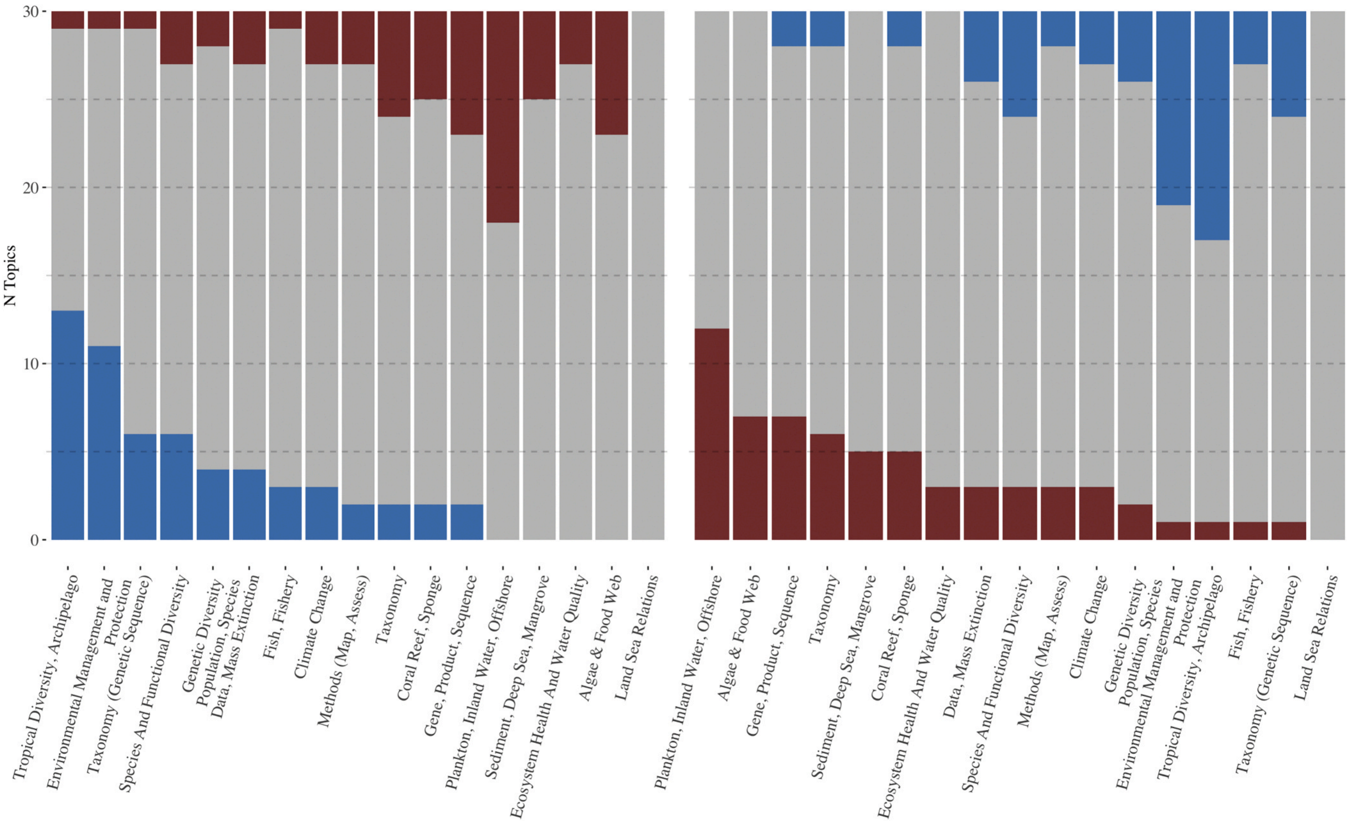

Fig. 5. Topics with high prevalence in collaborative/non-collaborative articles.

Table 2
Negative Binomial Regression Model Showing the Effect of Expertise.

	Estimated Parameters [90% Credibility Interval]
Intercept	-14.51 [19.46, −10.06]
log(Region A Output)	0.24 [−0.04, 0.55]
log(Region B Output)	1.22 [0.95, 1.53].
ΔExpertize	-0.05 [−0.05, −0.04]
Varying Intercepts	
Cluster Level (SD)	0.16 [0.11, 0.25]
Pair Level (SD)	0.51 [0.34, 0.79]
N	510

7. Discussion and concluding remarks

Marine biodiversity is a highly complex scientific field that is tightly connected to biodiversity research and ocean science. The current paper proposed to view this as a separate research field and attempted to delineate it using bibliometric data and state-of-the-art text analysis methodology. The current paper shows that marine biodiversity has both the characteristics of a homogeneous research field and several specificities that are distributed in both spatial and temporal dimensions. We study the dynamics and effects of scientific collaboration in the marine biodiversity field based on scientific text. We look at the emerging research field of marine biodiversity through the lens of “world-building” as defined by (Callon et al., 1986).

According to (Callon et al., 1986), scientists that work in specialized fields, through their association with a larger discipline (marine biodiversity in this case) are creating a field that is greater than the sum of its parts, and that has transcended the purely scientific dimension. All papers in our sample have a direct connection to ‘marine biodiversity’ through keywords (as per the design of this study); scientists working on very different topics have made either an explicit or an implicit decision to associate these topics to marine biodiversity. This willingness to associate oneself with a specific area of research, in our view, also proves that marine biodiversity is a rapidly emerging and separate research field from biodiversity and ocean science. It saw a significant increase in publication output,[fn1] that surpassed in magnitude the general inflation of scientific output. This rise is not only evident from the raw count of papers with “marine biodiversity” keywords, but also in certain cluster prevalence in the academic papers’ abstracts. For example, the *Environmental Management* cluster saw a significant rise in the beginning to mid-2010s in North America, Europe, and especially Oceania.

Several things could potentially explain this increasing interest in marine biodiversity as a whole field and certain clusters in particular from a scientific community. This rise coincides with several large-scale international programs. The Census for Marine Life, one of the biggest such initiatives, was a 10-year long global initiative designed specifically to assess and understand the biodiversity in Earth’s oceans (Snelgrove, 2010). This program ran from 2000 with the comprehensive results released in 2010, which could be connected to the increased interest in the field from the scientific community.

We show that the subfields of marine biodiversity (i.e., clusters) are not stable over time, that its clusters either gain or lose popularity in different regions with different rates. This is most evident with the aforementioned *Environmental Management* cluster, which overtook cluster prevalence in Oceania, Europe, and North America in the early to mid-2000s, but is also observable, for example, in Asia, where *Genetic Diversity* cluster has been steadily losing its popularity from the early 2010s, in favor of the *Environmental Management* cluster. These result show that marine biodiversity is a dynamic and evolving field, which still has not settled in its evolution.

---

[1] The average annual growth rate of the marine biodiversity field (WoS Core Collection data) is 15.77%, compared to an average of 4.55% for the total WoS Core Collection repository (years 1990–2018).

Since marine biodiversity is a very technologically-dependent and data-intense field, scientific and technological advances are also linked to the increased interest in the field. Specifically, high-tech innovations such as the DNA barcoding method (Hofmann and Gaines, 2008; Savolainen et al., 2005) that gained prominence in the mid-2000s allowed scientists to create a universal system for biodiversity classification and assign genetic lineage to previously unknown species. DNA barcoding is extensively used in marine biodiversity and is considered one of the major breakthroughs of the field in the XXI century (Radulovici et al., 2010; Trivedi et al., 2016).

Furthermore, around the same time, marine biodiversity has become deeply entrenched in the political dimension. As a "unifying imperative" the term was coined to designate the need for international regulations regarding the protection and sustainable use of biodiversity and the access and benefit sharing of its genetic properties (Vadrot, 2014). The national regulations that were put in place to comply with the Nagoya Protocol on Access and Benefit Sharing (Anon, 2015) have a significant impact on the research practices regarding biodiversity research, as they regulate access to specific sites of data collection and sharing of benefits that may arise from the utilization of MGRs collected in national waters (Blasiak et al., 2018). In order to clarify ownership of MGRs in the High Seas UN member states have started to negotiate a new treaty for the conservation and sustainable use of marine biodiversity beyond national jurisdiction (BBNJ). This interest in marine biodiversity from the political circles - seeking solutions to the degradation and commodification of marine resources- roughly coincides with the growth rate of the *Environmental Management* cluster in scientific abstracts, and while the current analytical setup of the paper and the analyzed data do not allow to make any causal claims, the correlation is too evident to be completely disregarded.

7.1. Selective world-building and regional specificities

Marine biodiversity is often considered a "big science". Our data suggest that while some aspects of marine biodiversity can indeed be thought of as a "big science", others – such as costal research and classical taxonomy – remain on a local level without exhibiting this character. Our results show that the situation is more complicated than a single global research field. The most prominent clusters across all regions are indeed the *Genetic Diversity* and the *Gene Sequencing* clusters. Nevertheless, our results show that under closer inspection, the "popularity" of clusters across different regions is rather uneven. Several regional differences are observed from data – the popularity of the *Environmental Management* cluster in the Global North (especially in Oceania, but also clearly present in Europe and North America) underscores the North/South difference in the thematic focus since this cluster is less represented in regions from the Global South, with Asian region's focus on genetic clusters (e.g., *Genetic Diversity*, *Gene Sequencing*, and *Genetic Taxonomy*), as well as South American interest in taxonomic clusters, both *Classical Taxonomy*, and *Genetic Taxonomy*.

These differences could be thought of as regional specificities, similar to what was theorized by (Vermeulen, 2013). Nevertheless, there are various possible explanations as to why these specificities arise. First, it is important to note that there are completely different interests associated with different aspects of marine biodiversity; as we discussed above, marine biodiversity includes a variety of different subfields that (apart from furthering the scientific knowledge) have secondary rationales. For example, the *Environmental Management* and *Climate Change* (although it is not very prominent in our dataset) clusters have very explicit political motivation behind them, with intergovernmental negotiations to create legally binding agreements and international organizations keeping a close watch on how international actors engage with compliance of international laws.

Other aspects of marine biodiversity science, like *Genetic Sequencing* has an additional commercial rationale, since marine genetic resources harvested from areas beyond national jurisdiction are of direct economic interest to both governments and corporations (Blasiak et al., 2018). Private corporations, for example, were both funding and reaping substantial monetary reward from their investments from the early days of biodiversity research (Blumenthal et al., 1986).

Furthermore, there are simply economic considerations of doing research (Kennedy, 1985) – especially one that requires significant financial infrastructure, like marine biodiversity. The topical focus, or a specialty of a region, can very much depend on the cost associated with said focus, and the willingness of the government to acquiesce to such cost – for example, some sources show that roughly 65% of all marine science research in South America was privately funded (Artigas and Escobar, 2001).

We can roughly bin various clusters into three different categories based on the cost associated with the research: relatively inexpensive, expensive but with economic benefit, and expensive without an immediate economic benefit. *Classical taxonomy*, for example, may not require the same financial influx as high-scale genetic sequence research; countries or regions that want to contribute to the marine biodiversity field, but lack a strong financial infrastructure may pursue less economically taxing research directions. On the other hand, regions may pursue an expensive, but economically lucrative approach, like, for example, *Genetic Sequencing* in order to extract monetary value from research. Finally, regions with a strong financial infrastructure have the luxury of engaging with high-cost, low economic benefit projects, such as evident with the Global North regions and their disproportionate prevalence of the *Environmental Management* cluster.

7.2. Selective world-building through scientific collaboration

Scientific collaboration within the "big science" fields is not purely an academic exercise, but also calculated political maneuvering (Hallonsten, 2014). Collaboration, is not a neutral act but is a practice of connecting the larger field, and as we show in our paper, is very dependent on the topic, having an ability to actively shape these topics. Importantly, preliminary results from this paper show that collaboration is directed, i.e., one would seek out a collaboration partner based on their experience in the subfield. Scientists collaborate with other scientists that possess something of value (be it expertise, access to data, etc.). Moreover, in some cases, a collaboration between scientists in biodiversity research is not only politically motivated but encouraged (e.g. in the context of the implementation of the Nagoya protocol).

The current paper extensively looked into the collaborative patterns of different marine biodiversity subfields and regions. We show that collaboration is quite dependent on a particular thematic cluster, with clusters positioned on a spectrum of higher/lower collaboration rates. Different collaboration rates for different (sub-)disciplines are not particularly surprising. For example, (Davidson Frame and Carpenter, 1979) theorized that different disciplines have varying collaboration rates – basic research has a higher collaboration rate than, for example, applied research. Tentatively, we can group clusters that are shaped by higher collaboration rates as having a more general/global interest – *Environmental Management*, *Genetic Taxonomy*, *Species and Functional Diversity*, *Genetic Diversity*, etc. Higher collaboration rates in the *Environmental Management* cluster could be related to the current political environment such as the BBNJ negotiations, while collaboration on *Genetic Taxonomy, Genetic Diversity, Functional Diversity* could be related to global initiatives like the Census of Marine Life. Another reason for higher collaboration rates in these clusters may be the fact that they are reliant on secondary data analysis (e.g., DNA barcoding), and/or statistical simulation models, which makes collaboration easier by removing the necessity to always be "on-site".

Conversely, clusters that are being less collaborated upon are the ones that either have a very regional character – such as *Inland/Offshore Waters, Food-web*, or *Classical Taxonomy*, where one needs to be "in the field" to perform research. In some case, like that of *Classical Taxonomy*, the research can be viewed as a "national endeavor" and (at least in

South America, where classical taxonomy has a very prominent position) an attempt to move away from the colonially imported notions of natural science (Salazar-Vallejo et al., 2007).

Thematic clusters that have some economic edge, like *Gene/Product Development* also have a relatively low collaboration rate. Furthermore, returning to the basic/applied research dichotomy outlined above, we can see that a similar pattern holds - basic research in marine biodiversity does indeed have more of a collaborative character than applied research.

We further looked into the way collaboration patterns are influenced by the region's expertize. We show that (controlling for the Thematic Cluster, Regional Scientific Output, and Region Pairs), there is a higher chance of collaboration on a cluster if a co-author (i.e., not the first author) has more expertize in the said cluster. For example, if Oceania has very high expertize in the *Environmental Management* cluster than Asia, there is a higher chance that in a region pair Asia-Oceania there is a higher collaboration rate on the *Environmental Management* cluster than in the Oceania-Asia pair.

7.3. Moving beyond text and future research needs

There are, however, several limitations associated with the current study. The current paper utilizes academic abstracts to map the field of marine biodiversity. While such an analysis is indispensable in order to continue to investigate the links between a scientific field and politics (or the political dimension of a scientific field), it remains difficult to establish direct connections with such a design.

Another limitation is the fact that the data were binned into continents, rather than individual countries. This decision was made out of several methodological considerations. In order for a topic model to produce stable results, the model must have enough data to estimate the effect of covariates (continents in this case) on the language, which is not the case for many individual countries. Additionally, some analyses in this paper require pairwise comparisons between continents, this would be computationally intractable with individual countries. If future research is interested in a similar approach to what we did in the current paper but wants to look into country-to-country comparisons, it should strategically select countries in order to make such an analysis feasible.

All articles in the current sample are by definition "marine biodiversity" articles, because we have pragmatically designed the study this way. This allows for an overarching analysis of the field using a large body of text, however leaves some questions unanswered. For example, what are the cluster boundaries? A cluster defined as "Fisheries" should not end on "marine biodiversity", but encompass a wider body of knowledge like freshwater biodiversity, economy, sustainability, etc. This whole discussion is, unfortunately, left out of our analysis, and would require a different approach and a different set of methods to engage with.

Nevertheless, we believe that the current paper provides a broad view on marine biodiversity as a field and regional collaboration patterns within it, and serves as an important first step in analyzing the field and its political dimensions. Future studies, armed with the overarching understanding of the general aspects of marine biodiversity should further investigate how the political dimension is affected by the topical output (and vice versa) or how marine biodiversity scientists experience and view the development of the field within which they work. This research is particularly timely now since the future of marine biodiversity is being negotiated at the United Nations level with a new treaty underway.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant No 804599). The ERC project MARIPOLDATA is led by Dr. Alice B.M. Vadrot (www.maripoldata.eu).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.envsci.2021.09.003.

References

AnonUnited Nations, 2015. Text And Annex of the Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization to the Convention on Biological Diversity. 1st ed. [ebook].Montreal: United Nations.

Artigas, C., Escobar, J., 2001. Perfil de la cooperacion para la investigacion cientifica marina en America Latina y el Caribe. CEPAL.

Blasiak, R., Jouffray, J.-B., Wabnitz, C.C., Sundström, E., Österblom, H., 2018. Corporate control and global governance of marine genetic resources. Sci. Adv. 4 (6), eaar5237.

Blumenthal, D., Gluck, M., Louis, K.S., Wise, D., 1986. Industrial support of university research in biotechnology. Science 231 (4735), 242–246.

Börner, K., Chen, C., Boyack, K.W., 2003. Visualizing knowledge domains. Annu. Rev. Inf. Sci. Technol. 37 (1), 179–255.

Bowker, G.C., 2000. Biodiversity datadiversity. Soc. Stud. Sci. 30 (5), 643–683.

Calero-Medina, C., Noyons, E.C., 2008. Combining mapping and citation network analysis for a better understanding of the scientific development: the case of the absorptive capacity field. J. Informetr. 2 (4), 272–279.

Callon, M., Rip, A., Law, J., 1986. Mapping the Dynamics of Science and Technology: Sociology of Science in the Real World. Springer.

Callon, M., Courtial, J.-P., Laville, F., 1991. Co-word analysis as a tool for describing the network of interactions between basic and technological research: the case of polymer chemsitry. Scientometrics 22 (1), 155–205.

Cobo, M.J., Lopez-Herrera, A.G., Herrera-Viedma, E., Herrera, F., 2011. An approach for detecting, quantifying, and visualizing the evolution of a research field: a practical application to the fuzzy sets theory field. J. Informetr. 5 (1), 146–166.

Cochrane, S.K., Andersen, J.H., Berg, T., Blanchet, H., Borja, A., Carstensen, J., Elliott, M., Hummel, H., Niquil, N., Renaud, P.E., 2016. What is marine biodiversity? towards common concepts and their implications for assessing biodiversity status. Front. Mar. Sci. 3, 248.

Coulter, N., Monarch, I., Konda, S., 1998. Software engineering as seen through its research literature: a study in co-word analysis. J. Am. Soc. Inf. Sci. 49 (13), 1206–1223.

Davidson Frame, J., Carpenter, M.P., 1979. International research collaboration. Soc. Stud. Sci. 9 (4), 481–497.

Hofmann, G.E., Gaines, S.D., 2008. New tools to meet new challenges: emerging technologies for managing marine ecosystems for resilience. Bioscience 58 (1), 43–52.

Kennedy, D., 1985. Government policies and the cost of doing research. Science 227 (4686), 480–484.

Latour, B., Knorr-Cetina, K., Mulkay, M., 1983. Science Observed: Perspectives on the Social Study of Science.

Leydesdorff, L., Nerghes, A., 2017. Co-word maps and topic modeling: a comparison using small and medium-sized corpora (n! 1,000). J. Assoc. Inf. Sci. Technol. 68 (4), 1024–1035.

Leydesdorff, L., Rafols, I., 2009. A global map of science based on the isi subject categories. J. Am. Soc. Inf. Sci. Technol. 60 (2), 348–362.

Litfin, K., 1994. Ozone Discourses: Science and Politics in Global Environmental Cooperation. Columbia University Press.

Noyons, E., Moed, H., Luwel, M., 1999a. Combining mapping and citation analysis for evaluative bibliometric purposes: a bibliometric study. J. Am. Soc. Inf. Sci. 50 (2), 115–131.

Noyons, E., Moed, H., Van Raan, A., 1999b. Integrating research performance analysis and science mapping. Scientometrics 46 (3), 591–604.

Partelow, S., Hornidge, A.-K., Senff, P., Stäbler, M., Schlüter, A., 2020. Tropical marine sciences: knowledge production in a web of path dependencies. PloS One 15 (2), e0228613.

Pendleton, L., Evans, K., Visbeck, M., 2020. Opinion: we need a global movement to transform ocean science for a better world. Proc. Natl. Acad. Sci. 117 (18), 9652–9655.

Pielke Jr., R.A., 2004. When scientists politicize science: making sense of controversy over the skeptical environmentalist. Environ. Sci. Policy 7 (5), 405–417.

Price, D.J., 1986. Little Science, Big Science. and Beyond. Columbia University Press New York.

Radulovici, A.E., Archambault, P., Dufresne, F., 2010. Dna barcodes for marine biodiversity: moving fast forward? Diversity 2 (4), 450–472.

Ramos-Rodríguez, A.-R., Ruíz-Navarro, J., 2004. Changes in the intellectual structure of strategic management research: a bibliometric study of the strategic management journal, 1980-2000. Strateg. Manag. J. 25 (10), 981–1004.

Roberts, M.E., Stewart, B.M., Tingley, D., et al., 2014a. stm: R package for structural topic models. J. Stat. Softw. 10 (2), 1–40.

Roberts, M.E., Stewart, B.M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S.K., Albertson, B., Rand, D.G., 2014b. Structural topic models for open-ended survey responses. Am. J. Political Sci. 58 (4), 1064–1082.

Rogers, A.D., Baco, A., Escobar-Briones, E., Currie, D., Gjerde, K., Gobin, J., Jaspars, M., Levin, L., Linse, K., Rabone, M., et al., 2021. Marine genetic resources in areas beyond national jurisdiction: promoting marine scientific research and enabling equitable benefit sharing. Front. Mar. Sci. 8, 600.

Salazar-Vallejo, S.I., Escobar-Briones, E., González, N.E., Suárez-Morales, E., Álvarez, F., de León-González, J.A., Hendrickx, M.E., 2007. Iniciativa mexicana en taxonomía: biota marina y costera. Cienc. Y. Mar. 11 (32), 69–77.

Sarewitz, D., 2004. How science makes environmental controversies worse. Environ. Sci. Policy 7 (5), 385–403.

Savolainen, V., Cowan, R.S., Vogler, A.P., Roderick, G.K., Lane, R., 2005. Towards writing the encyclopaedia of life: an introduction to dna barcoding. Philos. Trans. R. Soc. B: Biol. Sci. 360 (1462), 1805–1811.

Sierra-Correa, K., Déniz-González, W., Kyewalyanga, N., Santana-Casiano, S., Isensee, Y., et al., 2020. Research capacity and infrastructure. In: Isensee, k. (Ed.), Ioc-unesco, Global Ocean Science Report 2020-Charting Capacity for Ocean Sustainability. UNESCO Publishing, Paris, pp. 93–133.

Skolnikoff, E.B., 2001. The political role of scientific cooperation. Technol. Soc. 23 (3), 461–471.

Small, H., 2003. Paradigms, citations, and maps of science: a personal history. J. Am. Soc. Inf. Sci. Technol. 54 (5), 394–399.

Small, H.G., 1977. A co-citation model of a scientific specialty: a longitudinal study of collagen research. Soc. Stud. Sci. 7 (2), 139–166.

Tydecks, L., Jeschke, J.M., Wolf, M., Singer, G., Tockner, K., 2018. Spatial and topical imbalances in biodiversity research. PloS One 13 (7), e0199327.

United Nations, 2021. The Second Global Integrated Marine Assessment. World Ocean Assessment II. United Nations, New York.

Brondizio, E.S., Settele, J., Diaz, S., Ngo, H.T. 2019. Global assessment report on biodiversity and ecosystem services of the intergovernmental science-policy platform on biodiversity and ecosystem services.

Callon, M., Courtial, J.-P., Turner, W.A., Bauin, S., 1983. From translations to problematic networks: an introduction to co-word analysis. Inf. (Int. Soc. Sci. Counc.) 22 (2), 191–235.

Hallonsten, O., 2014. The politics of european collaboration in big science. In: In The Global Politics of Science and Technology, 2. Springer, pp. 31–46.

Snelgrove, P.V., et al., 2010. Discoveries of the Census of Marine Life. Cambridge University Press, Cambridge.

Takacs, D., 1996. Idea Biodivers.: Philos. Paradise.

Tolochko, P., Vadrot, A.B., 2021. The usual suspects? distribution of collaboration capital in marine biodiversity research. Mar. Policy 124, 104318.

Trivedi, S., Aloufi, A.A., Ansari, A.A., Ghosh, S.K., 2016. Role of dna barcoding in marine biodiversity assessment and conservation: an update. Saudi J. Biol. Sci. 23 (2), 161–171.

Upham, S., Small, H., 2010. Emerging research fronts in science and technology: patterns of new knowledge development. Scientometrics 83 (1), 15–38.

Vadrot, A.B., 2014. The Politics of Knowledge and Global Biodiversity. Routledge.

Vadrot, A.B., 2020. Building authority and relevance in the early history of ipbes. Environ. Sci. Policy.

L. Valdes et al., Global Ocean Science Report: The Current Status of Ocean Science Around the World, 2017.

Vermeulen, N., 2013. From darwin to the census of marine life: marine biology as big science. PloS One 8 (1), e54284.

Wallach, H.M., Murray, I., Salakhutdinov, R., Mimno, D., 2009. Evaluation methods for topic models. Proc. 26th Annu. Int. Conf. Mach. Learn. 1105–1112.

Warren, C.R., 2007. Perspectives on thealien'versusnative'species debate: a critique of concepts, language and practice. Prog. Hum. Geogr. 31 (4), 427–446.

Wilson, E.O., 1985. The biological diversity crisis. BioScience 35 (11), 700–706.

Wilson, E.O., et al., 1988. The current state of biological diversity. Biodiversity 521 (1), 3–18.

Zitt, M., 2005. Facing diversity of science: a challenge for bibliometric indicators. Meas. Interdiscip. Res. Perspect. 3 (1), 38–49.