Cognitive diversity and the future of crises: an analysis of the topic space of the biological sciences

Pierre Benz[*], Vincent Larivière[**], Diego Kozlowski[**], Natsumi S. Shokida[**] and Carolina Pradier[**]

[*] *pierre.benz@umontreal.ca*
0000-0003-1672-7311
School of Library and Information Science, University of Montreal, Canada

[**] *vincent.lariviere@umontreal.ca; diego.kozlowski@umontreal.ca; natsumi.solange.shokida@umontreal.ca; carolina.pradier@umontreal.ca*
0000-0002-2733-0689; 0000-0002-5396-3471; 0009-0008-9198-3888; 0009-0007-5058-6352
School of Library and Information Science, University of Montreal, Canada

This paper proposes to address the relationship between cognitive diversity and research topics among biologists. It asks whether biologists who are 'open' to a greater variety of topics are also more prompt to tackle issues relative to current global crises, or if some key topics like climate change, biodiversity and global health are confined to rather institutionally hermetic disciplinary landscapes. To answer this question, we propose to map a space of topics as a combination of latent topic modeling and multiple correspondence analysis. Such a method allows us to relate topics with proprieties of both journals and authors. It also provides an empirically informed framework to operationalize the cognitive diversity of biologists with reference to the distribution of their most prevalent vocabulary the space of topics. Sample for analysis is based on all publications (34,797) from all professors of biology in Switzerland between 2008 and 2020 (n=465).

1. Introduction

Recent times are confronted with crucial concerns about understanding and eventually resolving major crises, including global warming, biodiversity loss, and increasing risks of pandemics. Scholars and policy makers have been calling for more collaboration between science, economy, and politics for decades (McCright & Dunlap, 2000). The Covid-19 pandemic has been a salient example of the many barriers to such cooperation, with the devastating consequence of exacerbating natural and social crises (Gaudillière et al., 2021). These barriers certainly derive from structural differentiation between social spheres, whose agents do not pursue the same stakes depending on whether they occupy positions in the economic, political, or scientific spheres (Bourdieu, 1979). But these barriers also stem from crucial epistemological divergences. For instance, political support for solution to the environmental crisis is not so much a question of opposing 'pro-science' or 'anti-science' traditions as of differing intellectual traditions and moral conceptions of the notions of 'nature' and 'environment' (Bryant & Farrell, 2024). In the scientific field too, scholars struggle and even oppose in the ways they approach problems and propose solutions (Bourdieu, 1991). Among the natural sciences, the biological sciences are a cornerstone for addressing environmental, biodiversity, and global health issues (Alizon, 2020). These, however, do not stand for any cohesive epistemic macrocosm and have regularly come under scrutiny as highly hierarchical and heterogeneous.

One major concern in retracing the historical transformations of biology is the co-evolution of two distinct branches: functional biology and evolutionary biology (Wilson, 1994; Magner, 2002; Sapp, 2003; Rheinberger, 2010; Meloni, 2016). A seminal and enduringly debated article by the biologist Ernst Mayr (1961) delineates these two groups as follows: functional biologists ask *how* and focus on proximal causes to the explanation of life at the microbiological level, while evolutionary biologists ask *why* and focus on historical explanations at the population level. More recent scholarship has suggested that the biological sciences have overcome this fundamental cognitive opposition (Laland et al., 2011) and is calling for a 'reconciliation' of the discipline (Morange, 2011; Alizon and Méthot, 2018). Barberousse et al. (2009) stress that the need for articulating developmental with evolutionary biology, 'that will perhaps change our conception of evolution, is beginning to make itself felt in a most urgent way' (p.7). Many aspects are evoked, among which that 'Novelty in today's biology is coming from the realization that some concepts which are widely used in different or even disconnected contexts, but are still vague and in need of clarification, may yet be given reinforced methodological roles when clarified.' A focus on epistemic issues through for instance the concepts of 'organism', 'interactionism', 'adaptation', and 'gene' is thus considered as the 'royal road' for enhancing the future of biology.

Contrary to the hope for greater cognitive diversity in biology, studies in the history and the sociology of science have emphasized the persistence of strong hierarchical divisions that derive from the highly differentiated symbolic value accorded to functional and evolutionary biologists (Strasser, 2006; Morange, 2020; Larregue et al., 2020; Benz & Bühlmann, 2024). From the mid-20th onwards, functional biologists have benefited from major support from both science policy makers and funding agencies, as well as major pharmaceutical and biotechnology companies. Furthermore, they have been considered as more 'rational' than evolutionary biologists, which have remained in a dominated position, comparatively (Gros, 1993; Strasser, 2006; Gingras, 2012). In turn, molecular biologists have been criticized for their 'reductionism' (Wilson, 1994; Morange, 2020). Alongside certain biologists, scholars in the philosophy of science aim to reaffirm that epistemologies in biology are more 'assemblages' than institutionalized configurations (Rheinberger, 2010). There is no doubt that the realm of the biological sciences has extended far beyond the boundaries of the discipline. The integration of numerous extra-disciplinary concerns, concomitantly to the process of 'biologization' of engineering, chemistry, and medicine among others, occured 'occasionally to the chagrin of biologists, whose genuine interest in questions of development and reproduction was not shared by everybody keen to exploit biological knowledge' (Gugerli et al., 2010, p.347).

However, one must prevent from amalgamate the multidisciplinary nature of the 'life sciences' landscape and biology as a discipline. The unprecedented rise of the life sciences from the 1970s onward has been interpreted as a seminal illustration of an 'end' of the discipline and have supported the development of numerous policies to encourage interdisciplinarity (Gibbons et al., 1994; Ledford, 2015). Yet, scholars have demonstrated the resilience of disciplines, which still serve as organizational units in the current academic landscape (Whitley, 1986; Abbott, 2001; Shinn, 2002; Heilbron & Gingras, 2015). Addressing interdisciplinarity in biology, scholars have extensively focused on the relationships to more or less distant disciplines (Chen et al., 2015; Larivière et al., 2015). This literature provides our understanding of the discipline with meaningful insights, among which the fact that close interdisciplinarity is less valuable than long-distance interdisciplinarity. In addition, the benefits of interdisciplinary collaboration in terms of impact vary according to the intensity of citations specific to each discipline, hence a citation intensive discipline like biomedical research will associate an increase in interdisciplinarity to a decrease in citations (Larivière & Gingras, 2010). Beyond cognitive barriers to intradisciplinary reconciliation, one may thus face another barrier: that of the symbolic and institutional hierarchies of (sub-)disciplines. Recently, scholars in bibliometrics have been discussing the pros and cons of paper-level analyses (Shu et al., 2018; Larivière & Sugimoto, 2018) in the context of the increasing use of various topic modeling techniques (Kozlowski et al., 2021). What remains rather unexplored is the relationship between cognitive diversity and research topics among biologists. Therefore, we ask whether cognitively diverse biologists are more prompt to tackle issues relative to current global crises, or if some key topics like climate change, biodiversity and global health—as well as specific concepts, such as 'organism', 'interactionism', 'adaptation', and 'gene'—are confined to institutionally rather hermetic disciplinary universes.

2. Research strategy

We focus on a small but comprehensively documented sample of all Swiss professors of biology (n=465) in Switzerland between 2008 and 2020, along with all their publications as retrieved from the Web of Science database (n=34,797). To address the extent to which cognitive diversity relates to specific topics, we map a space of topics by combining latent Dirichlet allocation (LDA) (Blei et al., 2003; Blei, 2012; DiMaggio et al., 2013; McFarland et al., 2013) with multiple correspondence analysis (MCA) (Le Roux & Rouanet, 2004). This paper aligns closely with the ongoing development of the 'topicspace' package for R by Benz and Larsen (2024). For previous work using this method, see for instance Kropp and Larsen (2023), Rossier et al. (2023) and Benz et al. (2024). Mapping topics in the form of a geometrical space allows us to retrieve coordinates for both publications and their authors, which makes it possible to operationalize cognitive diversity as Euclidean distances. To associate given variables, i.e., journal discipline, with specific vocabulary, we calculate risk ratios (RR) as the relative risk of a certain event—being attributed a term—occurring in one group, such as all publications from a specific journal discipline, compared to the risk of it occurring in another group (the other publications). Hence, a greater distance between terms with the highest risk ratios indicates a higher level of cognitive diversity in an author. The results section proceeds as follows: *first*, we define and map the space of biology topics from the abstracts and titles of the publications. *Second*, we assess the degree to which the distribution of topics remains structured by significant institutional distinctions. We compare between 'biochemistry and molecular biology' and 'ecology' specialties as representative of respectively functional and evolutionary biologists. *Third*, we measure the degree of cognitive diversity of the professors by quantifying the distance between their respective research topics. Finally, we compare the distribution of these distances across disciplines, specific concepts as suggested by Barberousse et al. (2009), and terms related to current crises, namely 'climate', 'warming', 'biodiversity', 'covid', 'pandemics', and 'health'.

3. Results

The space of topics on which our analysis rely is as follows. After defining the topics using LDA, the most contributing topics are mapped into a factorial plan through their co-occurrence across publication abstracts. Each topic is associated with coordinates on each axis (or dimension) and a contribution to the formation of axes. Although Benzécri's modified rates suggest retaining three axes for the interpretation of the space' structure (Le Roux & Rouanet, 2004), we limit our analysis to the first two principal axes, which together account for 71.2% of the explained variance (table 1). Figure 1 displays the map of topics on the two principal axes.

Table 1. Variance and modified rates.

Dimension	**Axis 1**	**Axis 2**	**Axis 3**
Eigenvalue	0.06	0.03	0.03
Variance	5.6	3.0	2.8
Modified rates (%)	59.6	11.6	9.0
Cum. modified rates (%)	59.6	71.2	80.2

Figure 1: The space of biology topics.

The first dimension along the horizontal axis displays an opposition between topics associated to functional biology on the left, and evolutionary on the right, hence illustrating the historical divide of the biological sciences between two poles (Mayr, 1961). On the left, the most contributing topics 'cell.stem.line', 'signal.pathway.kinase', and 'protein.family.functional' contrast with those on the right, namely 'species.habitat.environmental', 'temperature.forest.tree', and 'population.genetic.trait'. It is noteworthy that topics related to population genetics are displayed on this pole. Indeed, as part of the modern synthesis (Wilson, 1994; Rheinberger, 2010), they directly associate with a 'evolutionary' approach to genetics. However, nuance is needed, while topics related to genetics and genomics is structuring the third dimension of the topic space. The key here is that the first axis represents the internal structure of biology, structurally opposing functional to evolutionary biologists.

The second dimension along the vertical axis is that of the relative autonomy of the discipline. On the top, we find topics related to the medical sciences, such as 'mouse.cell.deficient', 'disease.model.loss', and 'patient.control.clinical'. At the bottom of the space are topics related to chemical and physical approaches to life, such as 'structure.channel.structural', 'domain.bind.terminal', and 'membrane.lipid.chain'. While the first dimension is that of an internal structure, the second dimension reflects the degree of porosity of biology's boundaries with neighbouring disciplines. The next step is to project journals' disciplines into the space. Here, we calculated the risk ratio (RR) associated with all publications in biochemistry and molecular biology journals (figure 2) and ecology journals (figure 3) as referring to the Web of Science 'specialty' variable.

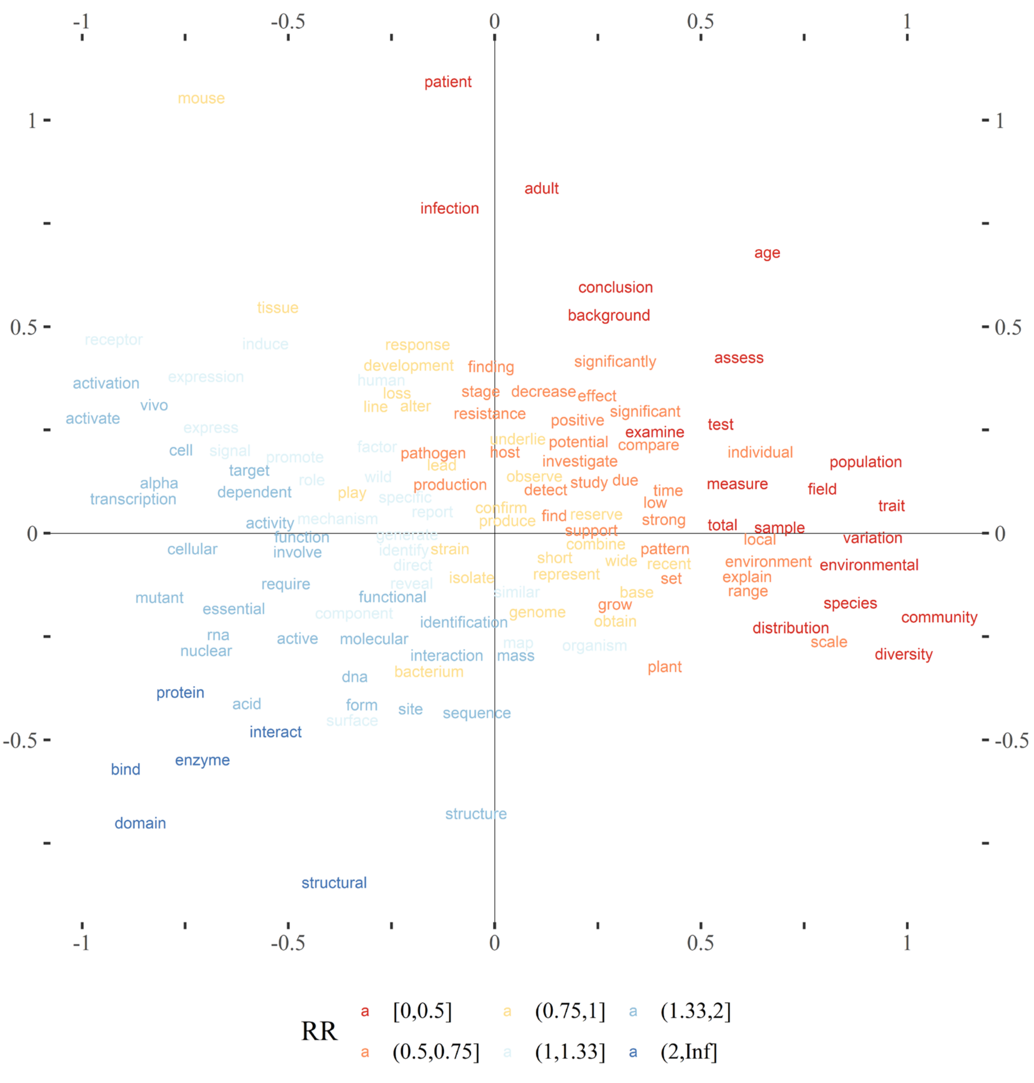

Figure 2: Risk ratios of terms according to the journal discipline: biochemistry and molecular biology.

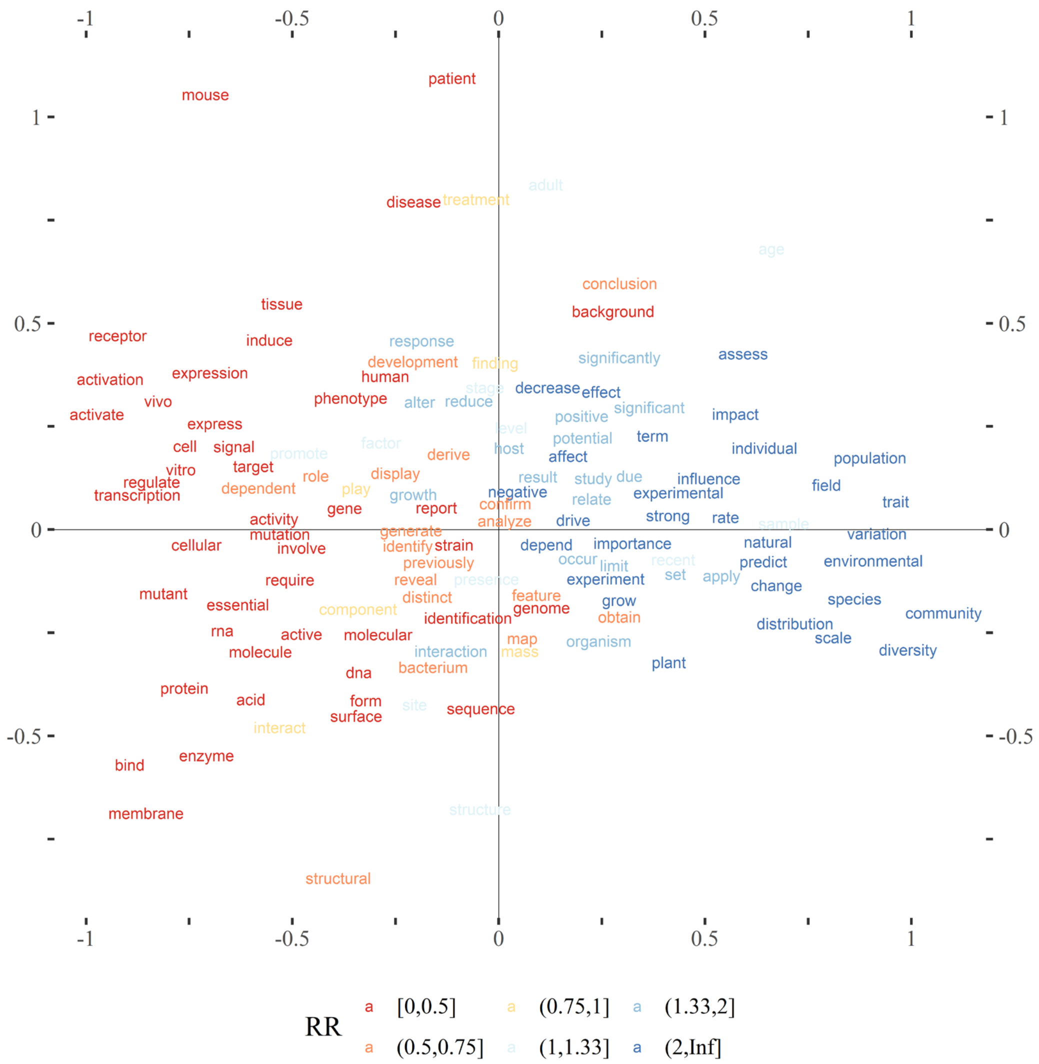

Figure 3: Risk ratios of terms according to the journal discipline: ecology.

What we observe here is unequivocal: publications in biochemistry and molecular biology journals locate on the left-side, while those in ecology journals locate on the right-side of the topic space. The clear association between the cognitive and the institutional structure of the discipline depicts a strong effect of (sub-)disciplinary frameworks in structuring the topics in biology.

Although the topic space is highly structured by disciplinary frameworks, it is entirely possible and probable that some authors will simultaneously mobilize topics located in different quadrants of the factorial plane. The last part of this preliminary results section is dedicated to assessing the degree to which such cognitive diversity associate with work on topics related to crises. We measure the degree of publication diversity of each author though k-means clustering on the publication coordinates. We then calculated the optimal number of clusters – the one that best represents the spatialization of publications – and the distances between the barycentre of these clusters on the x-axis (internal differentiation) and the y-axis (degree of porosity to neighbouring disciplines). Thus, an author whose publication coordinates are close will be committed to limited cognitive diversity. Conversely, an author whose publications form hubs in different locations will be deemed cognitively diverse. Moreover, it is possible to qualify this ‘openness’ according to the axis concerned, i.e., internal (axis 1) or external (axis 2).

As a final step, we compare the distribution of these distances along the first and the second axes. Here we defined the discipline of professors by the keywords they self-assign when submitting projects to the Swiss National Science Foundation (again, by comparing between molecular biology and ecology). We also compare between authors who are referring to specific terms in their publication' abstracts, namely 'climate', 'warming', 'biodiversity', 'covid', 'pandemics', and 'health'. We acknowledge that these ways of operationalizing disciplines and thematic focuses carry some limitation. Figure 4 displays the distribution of distances along the first (internal) and the second (external) axis.

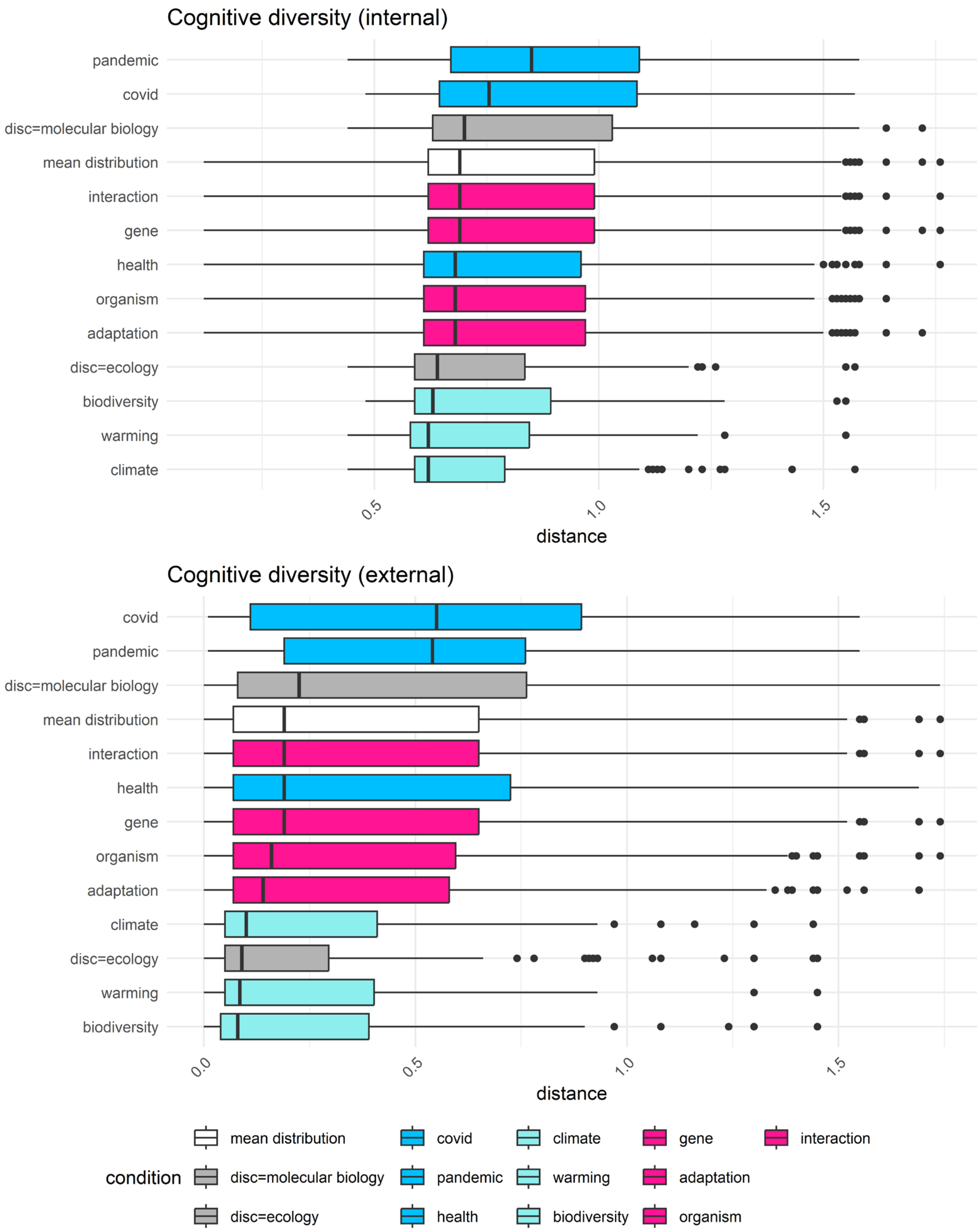

Figure 4: Distribution of the distances between the topics of professors.

Disciplines are displayed in grey. The four concepts from Barberousse et al. (2009) are displayed in pink. Other topics related to global health are displayed in blue, while those relating to environmental issues are displayed in green.

A first observation is that molecular biologists are much more prone to cognitive diversity than ecologists. One could argue for this is a consequence of the 'molecularization' of the natural sciences (Gugerli et al., 2010; Morange, 2020). However, our findings suggest that molecular biologists are also cognitively mobile *internally* within biology that is highly structured both cognitively and institutionally. This contrasts with the literature that depicts molecular biologists as 'reductionists', while ecologists would rather be described as practising a more 'open', more transversal, even more global biology (Gros, 1993; Wilson, 1994; Strasser, 2006; Rheinberger, 2010; Meloni, 2016).

A second observation concerning topics is that the four key concepts invoked by Barberousse et al. (2009) are used by authors with a cognitive diversity that is close to the average on both axes. For now, we do not find any relevant information concerning a specific use of these notions. What is striking, however, it the difference between topics linked to global health, and climate and environmental issues. Here, we clearly observe that especially publications mentioning 'pandemic' and 'covid' in their abstract are authored by professors with the highest degree of both internal and external cognitive diversity. Hence, it seems from this result that the opposition between functional microbiologists and evolutionary biologists in not as strong as expected by the literature (Alizon, 2020; Gaudillière et al., 2021). On the other side of the distribution, topics related to environment such as 'climate', 'warming', and 'biodiversity' are driven by professors with comparatively little cognitive diversity.

4. Discussion and conclusions

This proposition aimed to address the relationship between cognitive diversity and research topics among biologists, asking whether certain issues related to the current crises were supported by especially 'open' scholars. Preliminary findings stem from a prospective study to experience the feasibility of theoretically and methodologically approaching scholars' cognitive diversity through the variety of their research topics. While scholarship has deeply focused on almost every aspect of interdisciplinarity defined by institutional classifications, we believe that further exploration of knowledge diversity must engage in articulating cognitive structures with the relative autonomy of disciplines. In this regard, combining LDA with MCA seems promising and opens for many avenues. In particular, the proposed method can integrate variables concerning both publications and authors, e.g., gender, institutional affiliation, or prestige. The sociological framework can be well deserved here, as it would be possible to superpose the topic space with the space of individuals' endowment in various resources. Data can, and are likely to, be extended to a wider range, for instance to all biological publications listed on the science web. Finally, our preliminary findings call for more research into the relationship between research objects and disciplines, and the reasons why some topics are strongly associated with certain disciplines—climate, warming and biodiversity in our study case—while others are, on the contrary, 'undisciplined' (Larivière et al., 2018).

5. Bibliographic references

Abbott, A. (2001). *Chaos of Disciplines*. Chicago: The University of Chicago Press.

Alizon, S., & Méthot, P.-O. (2018). Reconciling Pasteur and Darwin to control infectious diseases. *PLoS Biol*, 16(1). e2003815. https://doi.org/10.1371/journal.pbio.2003815.

Alizon, S. (2020). *Evolution, Ecology, and Pandemics*. Paris: Points Sciences.

Barberousse, A., Morange, M. & Pradeu, T. (eds.) (2009). *Mapping the Future of Biology. Evolving Concepts and Theories*. Berlin: Springer.

Benz, P., & Bühlmann, F. (2024). Structural power and epistemologies in the scientific field: Why a rapid reconciliation between functional and evolutionary biology is unlikely. *Minerva*, 62(2), 229–251. https://doi.org/10.1007/s11024-023-09520-0.

Benz, P., & Larsen, A. G. (2024). topicspace: Mapping topic spaces combining latent dirichlet allocation with multiple correspondence analysis. R package version 0.1.0. https://github.com/BenzPierre/topicspace.

Benz, P., Kropp, K., Nobel, T. C., & Rossier, T. (2024). Homologies in fields of cultural production. Evidence from the European scientific field. *Poetics*, 107, 101945. https://doi.org/10.1016/j.poetic.2024.101945.

Blei, D. M, Ng, Y. N., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993–1022.

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77–84. http://dx.doi.org/10.1145/2133806.2133826.

Bourdieu, P. (1979). *La Distinction. Critique sociale du jugement*. Paris: Editions de Minuit.

Bourdieu, P. (1991). The peculiar history of scientific reason. *Sociological Forum*, 6(1), 3–26. https://doi.org/10.1007/BF01112725.

Bryant, J. C., & Farrell, J. (2024). Conservatism, the far right, and the environment. *Annual Review of Sociology*, 50, 273–296. https://doi.org/10.1146/annurev-soc-083023-035225.

Chen, S., Arsenault, C., Gingras, Y., & Larivière, V. (2015). Exploring the interdisciplinary evolution of a discipline: the case of Biochemistry and Molecular Biology. *Scientometrics*, 102(2), 1307–1323. https://doi.org/10.1007/s11192-014-1457-6.

DiMaggio, P., Nag, M., & Blei, D. (2013). Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of US government arts funding. *Poetics*, 41(6), 570–606. https://doi.org/10.1016/j.poetic.2013.08.004.

Gaudillière, J.-P., Izambert, C. & Juven, P.-A. (2021). *Pandémopolitique. Réinventer la santé en commun*. Paris: La Découverte.

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*. London: Sage Publications.

Gingras, Y. (2012). Le champ scientifique. In F. Lebaron & G. Mauger (eds.), *Lectures de Bourdieu* (pp.279–294). Paris: Ellipses.

Gros, F. (1993). *Regard sur la biologie contemporaine*. Paris: Gallimard.

Gugerli, D., Kupper, P., & Speich, D. (2010). *Die Zukunfts Machine. Konjonkturen der ETH Zürich*. Zurich: Chronos Verlag.

Heilbron, J., & Gingras, Y. (2015). La résilience des disciplines. *Actes de la Recherche en Sciences Sociales*, 210, 4–9. https://doi.org/10.3917/arss.210.0004.

Kozlowski, D., Dusdal, J., Pang, J., & Zilian, A. (2021). Semantic and relational spaces in science of science: Deep learning models for article vectorisation. *Scientometrics*, 126(7), 5881–5910. https://doi.org/10.1007/s11192-021-03984-1.

Kropp, K., & Larsen, A. G. (2023). Changing the topics: the social sciences in EU-funded research projects. *Comparative European Politics*, 21(2), 176–207. https://doi.org/10.1057/s41295-022-00313-5.

Laland, K. N., Sterelny, K., Odling-Smee, J., Hoppitt, W., & Uller, T. (2011). Cause and effect in biology revisited: is Mayr's proximate-ultimate dichotomy still useful? *Science*, 334(6062), 1512–1516.

Larivière, V., & Gingras, Y. (2010). On the relationship between interdisciplinarity and scientific impact. *Journal of the American Society for Information Science and Technology*, 61(1), 126–131.

Larivière, V., Haustein, S., & Börner, K. (2015). Long-distance interdisciplinarity leads to higher scientific impact. *PloS One*, 10(3), e0122565. https://doi.org/10.1371/journal.pone.0122565.

Larivière, V., & Sugimoto, C. R. (2018). *Mesurer la science*. Montréal: Les Presses de l'Université de Montréal.

Larivière, V., Plante, D. M., & Mongeon, P. (2018). Des objets disciplinés… aux indisciplinés. *Acfas Magazine*. https://www.acfas.ca/publications/magazine/2018/03/objets-disciplines-aux-indisciplines.

Larregue, J, Larivière, V. & Mongeon, P. (2020). On the institutional and intellectual division of labor in epigenetics research: A scientometric analysis. *Social Science Information*, 59, 117–143. https://doi.org/10.1177/0539018419898394.

Le Roux, B., & Rouanet, H. (2004). *Geometric Data Analysis: From Correspondence Analysis to Structured Data Analysis*. Springer Science & Business Media.

Ledford, H. (2015). How to solve the world's biggest problems. *Nature*, 525, 308–311. https://doi.org/10.1038/525308a.

Magner, L. N. (2002). *A History of the Life Sciences, Revised and Expanded*. Boca Raton: CRC Press.

Mayr, E. (1961). Cause and effect in biology. *Science*, 134(3489), 1501–1506.

McCright, A. M., & Dunlap, R. E. (2000). Challenging global warming as a social problem: An analysis of the conservative movement's counter-claims. *Social Problems*, 47(4), 499–522. https://doi.org/10.2307/3097132.

McFarland, D. A., Ramage, D., Chuang, J., Heer, J., Manning, C. D., & Jurafsky, D. (2013). Differentiating language usage through topic models. *Poetics*, 41(6), 607–625. https://doi.org/10.1016/j.poetic.2013.06.004.s

Meloni, M. (2016). *Political Biology: Science and Social Values in Human Heredity from Eugenics to Epigenetics*. London: Palgrave Macmillan.

Morange, M. (2011). Darwinisme et biologie moléculaire. In T. Heams, P. Huneman, G. Lecointre, & M. Silberstein (eds.) *Les mondes darwiniens: L'évolution de l'évolution* (pp.381–394). Paris: Éditions Matériologiques.

Morange, M. (2020). *The Black Box of Biology*. Cambridge: Harvard University Press.

Rheinberger, H.-J. (2010). *An Epistemology of the Concrete: Twentieth-Century Histories of Life*. Durham and London, Duke University Press.

Rossier, T., Benz, P., Larsen, A. G., & Kropp, K. (2023). The space of research topics in economics: Scientific position-takings and individual positions in Swiss economic science. *Œconomia. History, Methodology, Philosophy*, (13-2), 427–473. https://doi.org/10.4000/oeconomia.15359.

Sapp, J. (2003). *Genesis. The Evolution of Biology*. Oxford and New York: Oxford University Press.

Shinn, T. (2002). The triple helix and new production of knowledge: prepackaged thinking on science and technology. *Social Studies of Science*, 32(4), 599–614.

Shu, F., Julien, C. A., Zhang, L., Qiu, J., Zhang, J., & Larivière, V. (2019). Comparing journal and paper level classifications of science. *Journal of Informetrics*, *13*(1), 202–225. https://doi.org/10.1016/j.joi.2018.12.005.

Strasser, B. J. (2006). L*a fabrique d'une nouvelle science. La biologie moléculaire à l'âge atomique (1945–1964)*. Florence: Olschki.

Whitley, R. (1986). *The Intellectual and Social Organization of Science*. Oxford: Clarendon Press.

Wilson, Edward O. (1994). *Naturalist*. Covelo: Island Press.

Competing interests
Authors have no competing interests to declare.

Funding information
Swiss National Science Foundation (Grant No. 210805)
Fonds de recherche du Québec - Société et Culture through their support UNESCO Chairs
Social Sciences and Humanities Research Council of Canada (SSHRC).

Open science practices
This project relies on two main sources: the Web of Science, which is not openly available, and an original prosopographic database of Swiss biology professors. The latter is a compilation of a range of various sources, namely printed directories, activity reports of the Swiss universities and CV data, which were mandatory to retrieve their publications from the Web of Science database. Analyses were performed through the 'topicspace' R package, which development version is available: https://github.com/BenzPierre/topicspace.

Author contributions
Conceptualization: PB, VL, DK, NS, CP
Data curation: PB, DK
Formal analysis, Methodology, Software, Writing: PB
Funding acquisition: PB, VL