Multilateral benefit-sharing from digital sequence information will support both science and biodiversity conservation

Amber Hartman Scholz1✉, Jens Freitag2, Christopher H. C. Lyal3, Rodrigo Sara1,4, Martha Lucia Cepeda5, Ibon Cancio6, Scarlett Sett1, Andrew Lee Hufton1, Yemisrach Abebaw7, Kailash Bansal8, Halima Benbouza9, Hamadi Iddi Boga10, Sylvain Brisse11, Michael W. Bruford12, Hayley Clissold13, Guy Cochrane14, Jonathan A. Coddington15, Anne-Caroline Deletoille11, Felipe Garcia-Cardona16, Michelle Hamer17, Raquel Hurtado-Ortiz11, Douglas W. Miano18, David Nicholson13, Guilherme Oliveira19, Carlos Ospina Bravo20, Fabian Rohden21, Ole Seberg22, Gernot Segelbacher23, Yogesh Shouche24, Alejandra Sierra25, Ilene Karsch-Mizrachi26, Jessica da Silva17,27, Desreee M. Hautea28, Manuela da Silva29, Mutsuaki Suzuki30, Kassahun Tesfaye7,31, Christian Keambou Tiambo32, Krystal A. Tolley17,27, Rajeev Varshney33, María Mercedes Zambrano34 & Jörg Overmann1,35

Open access to sequence data is a cornerstone of biology and biodiversity research, but has created tension under the United Nations Convention on Biological Diversity (CBD). Policy decisions could compromise research and development, unless a practical multilateral solution is implemented.

Here, we lay out a framework for use of digital sequence information (DSI) that enables fair benefit-sharing, ensures open access to sequence data, strengthens biodiversity conservation and sustainable use, and leverages genomics and bioinformatics for international capacity-building. As Parties to the CBD meet again in-person in the coming months to negotiate the Global Biodiversity Framework, they must apply pragmatic, multilateral solutions to DSI that improve rather than impede global biodiversity targets.

The ability to decode and digitally archive DNA has revolutionized the life sciences and related fields. Sequence data, referred to as digital sequence information (DSI) in policy

A full list of author affiliations appears at the end of the paper.
circles, are key to scientific advancement and technological innovation in fields as diverse as medicine, food security, green energy production, and biodiversity conservation. For example, free and open access to the SARS-CoV-2 viral sequences enabled the rapid development of diagnostic kits and vaccines. Besides its relevance for the common good, DSI is also essential for many commercial applications and offers new perspectives for economic development worldwide. Nonetheless, there are serious concerns about the equitable distribution of these benefits from a global perspective.

Parties to the Convention on Biological Diversity (CBD, https://www.cbd.int/) have recognized that countries have a sovereign right to regulate access to their genetic resources by requiring users to obtain prior informed consent and accept mutually agreed terms that ensure benefit sharing, i.e., that a portion of the advantages or profits derived from use are shared with the providing country. This arrangement is often referred to as simply “ABS” for Access and Benefit-Sharing. These bilateral procedures between countries and users are codified in the CBD’s Nagoya Protocol, which puts forth transparent procedures, compliance mechanisms, and checkpoints that are intended to monitor and ensure benefit-sharing. While well-intentioned, this system has proven to be inefficient and often incurs high transaction costs to obtain the necessary permits.

The rules for accessing DSI produced from these same sovereign genetic resources are, however, generally unclear. DSI is typically held in online open-access databases, where use is disconnected from physical access and accompanying permits. Biodiverse nations, many of which are low- and middle-income countries (LMICs), believe their sovereign rights have been undermined because any potential monetary gains from DSI through commercialization are not shared back to them, as they would be with a genetic resource. Thus DSI is perceived as a loophole that inhibits fair and equitable benefit-sharing. Moreover, the political pressure to close this loophole is high. Multiple authors on this paper have attended online negotiations, in which negotiators have explicitly stated that international benefit-sharing from DSI is a precondition to reach political consensus on the Global Biodiversity Framework. In other words, DSI is a make or break issue for one of the most important decadal environmental deals.

Ironically, according to a recent high-level report, scientists must have open access to DSI to fulfill the aims of the Global Biodiversity Framework (https://www.un.org/sustainabledevelopment/blog/2021/07/a-new-global-framework-for-managing-nature-through-2030-1st-detailed-draft-agreement-debuts/) and the UN Sustainable Development Goals (https://sdgs.un.org/goals).

What are the current options?

Given these pressures, Parties want options. A range of monetary benefit-sharing options for DSI were recently synthesized by the CBD Secretariat. These range from a bilateral system similar to the CBD and its Nagoya Protocol, with access tightly coupled to benefit sharing, to multilateral mechanisms in which access is facilitated through standardized global rules for benefit-sharing. In March 2020, Laird et al. made a compelling case for multilateralism, and called for the scientific community to work together to develop policy options, but left unanswered what a policy framework could look like.

We respond to this call and provide a concrete framework for how DSI benefit-sharing could work (Fig. 1). The authors are members of the DSI Scientific Network, a group of scientists from different countries and economic settings that share convergent points of view in the DSI debate. We note that every country provides and uses DSI and, thus, everyone has something to lose and gain. Here, we map out a multilateral framework for DSI that will support biodiversity monitoring and conservation, maintain open access, improve the scientific record, share both non-monetary and monetary benefits, and enable green growth.

The DSI policy option that will not work

The core infrastructure for storing sequence data, the International Nucleotide Sequence Database Collaboration (INSDC) contains 228 million annotated sequences (1.5 billion reads, see https://www.ncbi.nlm.nih.gov/genbank/statistics/). The dataset is downloaded partially or completely 34 million times per year, used by more than 10–15 million unique users, and 99.5% of the 750 downstream sequence databases that pull and push DSI through the scientific ecosystem directly rely on the INSDC system. The data are linked to a further 1200 databases and hundreds of thousands of publications that, on average, cite 44 sequences per publication. Clearly, a bilateral system, modeled on the principles of the Nagoya Protocol, that required permission between the end-user and the country of origin for every sequence and user transaction, would be prohibitively complex, affect data interoperability, and be ill-suited for generating knowledge.

The DSI statistics above make clear that a human-based system would not work. Technological approaches to track and trace DSI usage would be expensive, complex, and could have a significant environmental footprint. Most importantly, a loss of open access would adversely impact the analytical capacities, data infrastructures, and academic systems in LMICs that rely on these open data infrastructures and make collaboration with researchers based in countries that enforce DSI use restrictions very unattractive.

A bilateral DSI system would also inadvertently create a competitive marketplace that would perversely incentivize the use of DSI from countries where no restrictions exist (e.g., USA, Germany) leading to DSI “forum shopping.” Ironically, DSI from LMICs that are needed most for global monitoring and protection of biodiversity would often be avoided which, in turn, would severely impede progress on the Global Biodiversity Framework.

A win-win multilateral option for DSI

A benefit-sharing framework for DSI clearly needs to be multilateral and should address five fundamental objectives:

1. **Open access.** Any future benefit-sharing system must guarantee open data access, which is required to be able to use and understand DSI. Only open access enables efficient and broad scale knowledge generation and capacity building. The existing core DSI infrastructure already fulfills these requirements.

2. **Simplicity.** The DSI data ecosystem is highly complex, even for expert users. For benefit-sharing to happen, the policy framework must be simple. If a complex regulatory layer is added on top of a complex technical system, it is doomed to fail.

3. **Harmonize.** The DSI dilemma is an opportunity to learn from current inefficiencies in ABS and minimize transaction costs. Furthermore, as DSI is being discussed in multiple international fora, DSI users need a harmonized framework to address benefit-sharing.

4. **Biodiversity.** Any mechanism needs to effectively support biodiversity conservation and sustainable use (the first two objectives of the CBD). The framework should incentivize and reward biodiversity knowledge generation, and fill in the blank spots on the world map of biodiversity.
Figure 1. Multilateral funding model for DSI benefit-sharing. Open access to DSI in databases (center blue box) is de-coupled from benefit-sharing (light green box). The multilateral fund (dark green box) can collect funds from micro-levies, voluntary contributions, innovative financing, or from patent royalties. Funding on a national level (dark green arrow) is weighted by the amount of sequences deposited from an individual LMIC, and together with the substantial non-monetary benefit-sharing from DSI (e.g., through international collaborations; lower gray arrow) enables capacity building (especially in the area of genomics and bioinformatics) as well as the build-up of a corresponding research infrastructure. Capacity building improves monitoring, conservation, and sustainable use of biodiversity (dark blue boxes), creating a positive feedback loop that leads to a continuous increase in knowledge generation on biodiversity. Simultaneously, greater availability of DSI creates additional positive feedback loops for generating non-monetary and monetary benefits (gray arrows). The asterisk (*) emphasizes that the use of DSI, and hence the benefits generated, largely depend on the use of DSI originating from several different sources (i.e., countries).

Increased responsibilities for scientists and databases

For this system to work, the scientific community needs to improve practice with regard to recording the country of origin for DSI and correct legacy problems. At present, only 16% of sequences in the INSDC have country of origin information associated with them; 44% of sequences without country data could and should have had country information provided by the submitting scientists. Geographical information is not only important for policymakers, but it is also useful scientific data that should be reported as a matter of scientific integrity and the FAIR principles, which place a strong emphasis on the importance of reliable “provenance” metadata. Importantly, the INSDC recently announced a new policy requiring spatio-temporal information demonstrating their commitment to scientific transparency and openness to change.

Challenges still remain. For example, sequences that are listed in a patent application often have country information listed in the patent application. Although these same sequences are automatically submitted to the INSDC by patent offices, the country information is not transferred. This remains a significant loss of transparency given the fact that patent-associated DSI represents 20% of the INSDC dataset.

Nevertheless, scientific databases are not regulatory entities. They are scientific infrastructures with a public mission. In our view, they can play a supportive role in the policy process by providing data to policymakers. As part of the Global Biodiversity Framework, DSI-holding databases could be called upon to coordinate and deliver “country reports”—global analysis of DSI submissions, user trends, capacity building and advancement on DSI provenance transparency (see above). These data could be linked back to the benefit-sharing calculation. Given that the INSDC is an American, European, and Japanese collaboration (all high-income countries and the US
is a CBD observer), it would likely be advisable to work in close collaboration with a more diverse group of databases. Here, the increasingly active Global Biodata Coalition could play a leading role (https://globalbiobiodata.org/).

There remain unanswered questions such as how to ensure fairness to indigenous people and local communities in a multilateral system. This could be encouraged via a new data tag for indigenous people and local communities, for example³ and simply making IPLCs directly eligible to receive monetary funds. It would also be counter-productive if users began to dump low-quality or repetitive DSI into the databases to “game” the system and artificially increase a country’s DSI proportion in the database. Here standards and internal quality checks would be essential and are actually realizable. Finally, economic modelling must be conducted to determine which of the monetary mechanisms listed above are most likely to be cost-effective, financially productive, and to minimize administrative costs.

DSI capacity-building enables “leapfrogging” towards the bioeconomy

While all countries use and produce DSI, there remain significant inequalities. Practical issues ranging from more expensive access to molecular biological reagents, slower internet bandwidth that limits high-throughput analyses, financial limitations for research funding, limited bioinformatics training and career development opportunities, as well as brain drain, routinely limit those of us working in LMICA. Thus, any DSI benefit-sharing framework must support technical capacity building focused on genomics and bioinformatics. The goal should be to facilitate a “leapfrog” effect in which LMIC scientists are trained to exploit DSI even while inequalities in high-tech sequencing or laboratory infrastructure including technology transfer are still being addressed. With advances in cloud computing, open-source software, and open access DSI databases, the gaps are easier to fill than ever before if strategic investments are made.

Science-focused capacity development within the CBD must be aimed both at conservation and building up the bioeconomy through sustainable use of bioresources. Local scientists and regional or national science academies should be involved in agenda-setting. Matchmaking platforms should be established that connect scientists across the globe and build up human capital enabling sustainable development.

Call to action

Decisions made at the Geneva meetings in March 2022 and, ultimately, at the 15th Conference of the Parties to the CBD later in 2022, will affect DSI practitioners for a generation to come. The crux of the decision will be whether governments want “control” over DSI through a bilateral system or whether they see the opportunity for greater scientific advancement and ultimately greater monetary value if they choose openness through a multilateral system.

Given the importance of DSI for the scientific community, we call for policymakers to affirm multilateralism at the COP15 meeting and de-couple access to DSI from benefit-sharing—to break the “ABS” acronym into “OA” and “BS” where open access, benefit-sharing, capacity building, and biodiversity conservation thrive. We encourage policymakers to engage with their national DSI practitioners and with our Network to work together towards a thoughtful, workable and informed compromise.

Received: 15 September 2021; Accepted: 28 January 2022; Published online: 23 February 2022

References

1. Schroeder, D. Benefit sharing: it’s time for a definition. J. Med. Ethics 33, 205–209 (2007).
2. Prathapan, K. D. et al. When the cure kills—CBD limits biodiversity research. *Science* 360, 1405–1406 (2018).
3. IPBES. Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. Zenodo https://doi.org/10.5281/zenodo.5513154 (2019).
4. Policy options for Access and Benefit Sharing and Digital Sequence Information. Secretariat of the Convention on Biological Diversity https://www.cbd.int/abs/DSI-webinar/DSIPolicyOptions2021.pdf (2021).
5. Laird, S. et al. Rethink the expansion of access and benefit sharing. *Science* 367, 1200–1202 (2020).
6. Scholz, A. H. et al. Myth-busting the provider-user relationship for digital sequence information. *GigaScience* 10, gia085 https://doi.org/10.1093/gigascience/gia085 (2021).
7. Rohden, F., Huang, S., Droge, G. & Scholz, A. H. Combined study on digital sequence information in public and private databases and traceability. Report No. CBD/DSI/AHTEG/2020/1/4 https://www.cbd.int/doc/elt/186/d793/57cb141ca40cb668e7cd9584/dsi-ah-teg-2020-01-04-en.pdf (2020).
8. Bockmann, F. A. et al. Brazil’s government attacks biodiversity. *Science* 360, 865–865 (2018).
9. Ribeiro, C. A. et al. How ownership rights over microorganisms affect infectious disease control and innovation: A root-cause analysis of barriers to data sharing as experienced by key stakeholders. *PLoS ONE* 13, e0195885 (2018).
10. Rourke, M., Eccleston-Turner, M., Phelan, A. & Gostin, L. Policy opportunities to enhance sharing for pandemic research. *Science* 368, 716–718 (2020).
11. Tydecks, L., Jeschke, J. M., Wolf, M., Singer, G. & Tockner, K. Spatial and topical imbalances in biodiversity research. *PLoS ONE* 13, e0199327 (2018).
12. Wilkinson, M. D. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* 3, 160018 (2016).
13. Liggins, L., Hudson, M. & Anderson, J. Creating space for Indigenous perspectives on access and benefit-sharing: Encouraging researcher use of the Local Contexts Notices. *Mol. Ecol.* 30, 2477–2482 (2021).

Acknowledgements

We thank the secretariat of the DSI Scientific Network, Isabelle Coche, Zaira Lanna, and Benjamin Robinson for their highly professional coordination of the network. This work was funded by the German Federal Ministry of Education and Research (BMBF) WiLDsI 031B0862 (A.H.S., J.F., and J.F.) and Horizon Europe EVA–GLOBAL 871029 (A.H.S.). I.K.M. was supported by the National Center for Biotechnology Information of the National Library of Medicine, National Institutes of Health.

Author contributions

Conceptualization: A.H.S., J.F., C.H.C.L., R.S., M.L.C., I.C., S.S., Y.A., K.B., H.B., H.I.B., S.Y., M.W.B., H.C., G.C., J.A.C., A.D., F.G.C.M., H.R.O., D.W.M., G.O., C.O.B., F.B., O.S., G.S., Y.S., A.S., J.d.S., M.d.S., M.S., K.T., K.A.T., M.M.Z., and J.O. Visualization: J.O., I.C., S.S., R.S., J.H.C.L., G.C., and A.H.S. Funding acquisition: A.H.S., J.F., and J.O. Writing—original draft: A.H.S., R.S., M.L.C., C.H.C.L., I.C., and S.S. Writing—review & editing: A.H.S., J.F., C.H.C.L., R.S., M.L.C., I.B., S.S., A.L.H., D.N., M.S., S.B., M.M.Z., O.S., K.T., K.A.T., R.H.O., J.d.S., C.K.T., R.V., J.O., D.H., and I.K.M.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Amber Hartman Scholz.

Peer review information *Nature Communications* thanks David Smith and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022

1Leibniz Institute DSMZ German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany. 2Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Seeland, Germany. 3Natural History Museum, London, UK. 4One Planet Solutions, Montpellier, France. 5Universidad de los Andes, Bogotá, Colombia. 6Plentzia Marine Station (PIE-UPV/EHU), European Marine Biological Resource Centre – Spain (EMBRC-Spain), Plentzia, Spain. 7Ethiopian Biotechnology Institute, Addis Ababa, Ethiopia. 8National Academy of Agricultural Science and Global Plant Council, New Delhi, India. 9National Council of Scientific Research and Technologies (NCSRT), Algiers, Algeria. 10Ministry of Agriculture, Livestock, Fisheries and Cooperatives, Nairobi, Kenya. 11Institut Pasteur, Paris, France. 12School of Biosciences, Cardiff University, Cardiff, UK. 13Wellcome Sanger Institute, Hinxton, UK. 14European Molecular Biology Laboratory European Bioinformatics Institute (EMBL-EBI), Hinxton, UK. 15Global Genome Initiative, Smithsonian National Museum of Natural History, Washington, DC, USA. 16Alexander von Humboldt Biological Resources Research Institute, Bogota, Colombia. 17South African National Biodiversity Institute, Cape Town, South Africa. 18University of Nairobi, Nairobi, Kenya. 19Instituto Tecnologico Vale (ITV), Belem, Brazil. 20Ministry of Environment and Sustainable Development, Bogota, Colombia. 21University of Lethbridge, Lethbridge, Canada. 22Natural History Museum of Denmark, Copenhagen, Denmark. 23University of Freiburg, Freiburg, Germany. 24National Centre for Cell Science, Pune, India. 25Mariano Galvez University, Guatemala City, Guatemala. 26National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD, USA. 27Centre for Ecological Genomics and Wildlife Conservation, University of Johannesburg, Johannesburg, South Africa. 28University of the Philippines Los Banos, Laguna, Philippines. 29Fundação Oswaldo Cruz (FIOCRUZ), Rio de Janeiro, Brazil. 30National Institute of Genetics, Mishima, Japan. 31Institute of Biotechnology, Addis Ababa University, Addis Ababa, Ethiopia. 32Centre for Tropical Livestock Genetics and Health (CTLGH) - International Livestock Research Institute (ILRI), Nairobi, Kenya. 33Murdoch University, Murdoch, Australia. 34Corporación CorpoGen, Bogotá, Colombia. 35Technical University of Braunschweig, Braunschweig, Germany.

✉email: amber.h.scholz@dsmz.de