



Convention on Biological Diversity

Distr.
GENERAL

CBD/DSI/AHTEG/2020/1/INF/1
4 February 2020

ENGLISH/SPANISH ONLY

AD HOC TECHNICAL EXPERT GROUP ON
DIGITAL SEQUENCE INFORMATION ON
GENETIC RESOURCES

Montreal, Canada, 17-20 March 2020

COMPILATION OF VIEWS AND INFORMATION ON DIGITAL SEQUENCE INFORMATION ON GENETIC RESOURCES SUBMITTED PURSUANT TO PARAGRAPHS 9 AND 10 OF DECISION 14/20

Note by the Executive Secretary

INTRODUCTION

1. The Conference of the Parties to the Convention on Biological Diversity adopted, at its fourteenth meeting, decision [14/20](#) on digital sequence information on genetic resources. In paragraph 9 of decision 14/20, the Conference of the Parties invited Parties, other Governments, indigenous peoples and local communities, relevant stakeholders and organizations to submit their views and information:

(a) To clarify the concept, including relevant terminology and scope, of digital sequence information on genetic resources and if and how domestic measures on access and benefit-sharing consider digital sequence information on genetic resources;

(b) On benefit-sharing arrangements from commercial and non-commercial use of digital sequence information on genetic resources.

2. The Conference of the Parties also invited Parties, other Governments and indigenous peoples and local communities to submit information on their capacity-building needs regarding the access, use, generation and analysis of digital sequence information on genetic resources, in particular for the three objectives of the Convention (decision 14/20, para. 10).

3. Further to the request of the Conference of the Parties specified in paragraph 11 (a) of decision 14/20, the Executive Secretary has compiled and synthesized the views and information submitted pursuant to paragraphs 9 and 10 of decision 14/20. Accordingly, the present document contains the compilation of the views and information while the synthesis of the views and information is made available as document CBD/DSI/AHTEG/2020/1/2.

4. This compilation reproduces the full text of all the submissions in the languages in which they were received by the Secretariat. Each submission has also been made available on the website of the Convention.¹

5. Submissions were received from the following:

(a) Parties to the Convention: Argentina, Australia, Belarus, Brazil, Canada, Colombia, Costa Rica, Ethiopia, European Union and its Member States, India, Iran, Japan, Madagascar, Mexico, Republic of Korea, South Africa, and Switzerland.

¹ <https://www.cbd.int/dsi-gr/2019-2020/submissions/>

- (b) A non-Party: United States of America;
- (c) Intergovernmental and other relevant organizations and stakeholders:
- African Union Commission (on behalf of the African Group);
 - Secretariat of the International Treaty on Plant Genetic Resources for Food and Agriculture;
 - Consortium of European Taxonomic Facilities (CETAF);
 - Consortium of German Natural History Collections (Deutsche Naturwissenschaftliche Forschungssammlungen);
 - German Life Sciences Association (Verband Biowissenschaften, Biologie und Biomedizin in Deutschland) and the Leibniz Biodiversity Research Alliance (Leibniz Verbund Biodiversität);
 - Chartered Institute of Patent Attorneys (CIPA);
 - EcoHealth Alliance;
 - EuropaBio - The European Association for Bioindustries, International Barcode of Life Consortium (iBOL);
 - International Chamber of Commerce (ICC);
 - International Fragrance Association (IFRA) and International Organization for the Flavor Industry (IOFI);
 - International Nucleotide Sequence Database Collaboration (INSDC);
 - Japan Bioindustry Association (JBA);
 - A joint statement from a group of stakeholder organizations,²
 - League of European Research Universities (LERU);

² These include : Signatory Organizations: (1) International: International Community of Breeders of Asexually Reproduced Alliance Ornamental and Fruit Plants, CropLife International, Genomic Standards Consortium, Global Genome Biodiversity Network, Invertebrate Global Genomics Alliance, Global Open Data for Agriculture & Nutrition, International Biocontrol Manufacturers Association, International Council of Biotechnology Associations, International Chamber of Commerce, International Federation of Pharmaceutical Manufacturers and Associations, The International Fragrance Association, International Organisation of the Flavor Industry, International Probiotics Association, International Seed Federation, Society for the Preservation of Natural History Collections, World Vegetable Center;

(2) Regional: African Seed Trade Association, Association of Manufacturers and Formulators of Enzyme Products, Asia & Pacific Seed Association, Confederation of European Yeast Producers (COFALEC), European Biostimulants Industry Council, European Cooperative Programme for Plant Genetic Resources (On behalf of 22 National Coordinators of ECPGR countries: Albania, Bosnia and Herzegovina, Bulgaria, Czech Republic, Estonia, Finland, Germany, Ireland, Italy, Latvia, Lithuania, Montenegro, The Netherlands, Portugal, Romania, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey), European Food & Feed Cultures Association, European Fermentation Group (CEFIC Sector Group), European Federation of Pharmaceutical Industries and Associations, European Seed Association, Enzyme Technical Association, European Association for Research in Plant Breeding, European Association for Specialty Yeast Products, European Association for Bioindustries, European Technology Platform (ETP), EU Specialty Food Ingredients, Specialty Feed Ingredients industry, FoodDrink Europe, Lactic Acid Bacteria Industrial Platform, Seed Association of the Americas;

(3) National: National Brazilian Association of Industry and Trade of Food Ingredients and Additives, L'Association des Fabricants de Compléments pour l'Alimentation Animale, American Institute of Biological Science, American Ornithological Society, American Society for Microbiology, Asociación Nacional de Obtentores Vegetales, American Phytopathological Society, Associazione Italiana Sementi, German Plant Breeders' Association, Biotechnology Innovation Organisation, Centre for Biodiversity Genomics - University of Guelph (CBG), Deutschen Industrievereinigung Biotechnologie, The Graduate School - Experimental Plant Sciences, Ethiopian Seed Association, German Association for Plant Breeding, GigaDB, GigaScience, Holland Bio, Leibniz-Institut für Pflanzengenetik und Kulturpflanzenforschung, Interpharma PH, Intellectual Property Owners Association, Japan Bioindustry Association, Japan Enzyme Association, Japan Pharmaceutical Manufacturers Association, Korea Biotechnology Industry Organization, Korea Cosmetic Association, Korea Health Supplements Association, Korea Pharmaceutical and Bio-Pharma Manufacturers Association, Korea Pharmaceutical Traders Association, Natural Science Collections Alliance, Personal Care Products Council, Plantum, Phycological Society of America, South African National Seed Organisation, Science Industries Switzerland, Society for Industrial Microbiology and Biotechnology, Italian Phytopathological Society, Swiss Biotech Association, French Specialty Food Ingredients Association, University of California Conservation Genomics Consortium, National Union of the Academies for the Sciences applied to Development of Agriculture, Food Security and Environment Protection (UNASA), University of Leuven, Union Française des Semenciers (UFS), US Culture Collection Network, US Department of Energy Joint Genome Institute Prokaryotic Super Program User Advisory Committee (members signing as individuals and not as employees/agents of the US government), Verband Biologie, biowissenschaft & biomedizin, Flanders Institute for Biotechnology

- A joint submission from the Natural History Museum UK, Royal Botanic Gardens Kew, and Royal Botanic Gardens Edinburgh;
- Natural History Museum Vienna;
- Society for the Preservation of Natural History Collections (SPNHC);
- Third World Network;
- UK BioIndustry Association (BIA); and
- Wellcome Sanger Institute.

INTRODUCTION	1
A. SUBMISSIONS FROM PARTIES	5
Argentina	5
Australia	6
Belarus	8
Brazil	10
Canada	18
Colombia	23
Costa Rica	25
Ethiopia	27
European Union and its Member States	28
India	30
Iran	32
Japan	33
Madagascar	34
Mexico	35
Republic of Korea	37
South Africa	38
Switzerland	40
B. SUBMISSION FROM A NON-PARTY	42
United States of America	42
C. SUBMISSIONS FROM INTERGOVERNMENTAL AND OTHER RELEVANT ORGANIZATIONS AND STAKEHOLDERS	43
African Union Commission on behalf of the African Group	43
Secretariat of the International Treaty on Plant Genetic Resources for Food and Agriculture (Plant Treaty)	49
Consortium of European Taxonomic Facilities (CETAF)	51
Consortium of German Natural History Collections (Deutsche Naturwissenschaftliche Forschungssammlungen), German Life Sciences Association (Verband Biowissenschaften, Biologie und Biomedizin in Deutschland) and the Leibniz Biodiversity Research Alliance (Leibniz Verbund Biodiversität)	56
Chartered Institute of Patent Attorneys (CIPA)	60
EcoHealth Alliance	64
EuropaBio - The European Association for Bioindustries	64
International Barcode of Life Consortium (iBOL)	66
International Chamber of Commerce (ICC)	68
International Fragrance Association (IFRA) and International Organization for the Flavor Industry (IOFI)	71
International Nucleotide Sequence Database Collaboration (INSDC)	72
Japan Bioindustry Association (JBA)	74
Joint statement by public and private sector organizations, academic and scientific institutions, data repositories and collections representing a broad range of stakeholders	78
League of European Research Universities (LERU)	80
Natural History Museum UK, Royal Botanic Gardens Kew, and Royal Botanic Gardens Edinburgh	82
The Natural History Museum Vienna (NHMW)	87
Society for the Preservation of Natural History Collections (SPNHC)	89
Third World Network	93
UK BioIndustry Association (BIA)	98
Wellcome Sanger Institute	103

A. SUBMISSIONS FROM PARTIES

Argentina

SUBMISSION

I. Terminología utilizada en este ámbito

1. A criterio de Argentina, de la interpretación de los términos de la expresión “**Información Digital sobre Secuencias de Recursos Genéticos**” surge que la “información”, como componente del “recurso genético”, está incluida en la definición de “**material genético**” del artículo 2 del Convenio.
2. Asimismo, en los trabajos preparatorios del Convenio de Biodiversidad (CBD), las Partes coincidían en que la referencia al acceso a la información y datos de los recursos genéticos era un elemento importante para la conservación y el uso sustentable, garantizando el beneficio mutuo por el uso de los recursos.
3. En consecuencia, la “*información*” y datos forman parte del concepto “*recurso genético*”, por lo que no es necesario realizar modificación alguna para su tratamiento.
4. Con relación a la cuestión terminológica, la terminología que mejor refleja el uso actual en el ámbito científico es “*información genética*”, que es preferible a “información digital sobre secuencias”. Ello porque
 - a) el término “*digital*” supondría una sola forma de transmisión de la información, pero pueden existir otras formas de transmisión de la información que no fueran digitales; y b) el término “*secuencia*” sólo refiere a la información contenida en la posición secuencial de nucleótidos en ADN o ARN y no contempla otra información contenida en las biomoléculas.

Antecedentes

Al analizar la cuestión de la secuenciación digital y su efecto sobre los objetivos de CBD, se genera la necesidad de clarificar la definición de los conceptos “recursos genéticos” o “material genético” para determinar si la información de las secuencias genéticas de los recursos constituye o no un recurso genético.

Definiciones

Por “**recursos genéticos**” se entiende el material genético de valor real o potencial.

Por “**material genético**” se entiende todo material de origen vegetal, animal, microbiano o de otro tipo que contenga unidades funcionales de la herencia.

Conceptualizaciones

En relación con la definición del concepto “**material genético**”, entre las varias acepciones posibles, el Diccionario de Oxford, al considerar la palabra “**material**” como sustantivo, la define como:

1- a) la materia de la cual una cosa está hecha o de la cual puede hacerse; b) las partes constitutivas de algo; y

2- información, evidencia, ideas, etc, que pueden usarse para crear un libro u otra obra. (en igual sentido Collins).

Si bien la palabra "material" refiere a "materia" que es una sustancia física y tangible, y que en general se distingue y opone a mente y espíritu (realidades intangibles), el término "material" no debe confundirse con materia. En este sentido, en el campo de los recursos genéticos, en general, la definición del término "material" admite la interpretación de que el mismo incluye la información asociada con el recurso genético, del cual esta información es parte constitutiva, sin importar el modo en que se transmite.

Por su parte, el término “**genético**” es un adjetivo que deriva de origen y el mismo diccionario la define como:

1- a) De o perteneciente a, o que involucra el origen; surgir de un origen común; b) de o perteneciente a la Genética o los genes; hereditario; y 2- generativo, productivo.

La “herencia genética” significa entre otras cosas:

1- lo que es transmitido de una generación a otra mediante los genes; 2- la constitución genética de un individuo. Remite al tema de la reproducción del material, la cual dado los avances científicos y tecnológicos actuales puede efectuarse a partir del recurso genético in situ, ex situ o por medio de la secuencia genética del recurso (sea esta digital o de otro formato, por ej. análogo).

En vista de las consideraciones anteriores, a criterio de Argentina, de la interpretación literal de los términos surge que la “información”, como componente del “recurso genético”, está incluida en la definición de “material genético” del artículo 2 del Convenio.

La cuestión de la “información” en los trabajos preparatorios del Convenio de Biodiversidad

Al momento de la negociación del Convenio de Biodiversidad, las Partes coincidían en que la referencia al acceso a la información y datos de los recursos genéticos era un elemento importante para la conservación y el uso sustentable (...) garantizando el beneficio mutuo por el uso de los recursos (Report of the ad hoc working group on the work of its second session in preparation for a legal instrument on biological diversity of the planet UNEP/BIO.DIV.2/3 - 23 FEBRUARY 1990).

Asimismo, cuando los negociadores discutieron el objetivo del Convenio definieron algunos consensos mínimos, tales como *the scope of access to biological diversity should include genetic material (...)* It also included direct physical access to biological diversity, as well as, indirect access, e.g. access to information. (Report of the ad hoc working group of legal and technical experts on biological diversity on the work of its second session UNEP/BIO.DIV/WG.2/2/5 - 7 MARCH 1991).

Estos consensos se pueden ver reflejados en la redacción de los primeros borradores, por ejemplo, el Artículo 13 sobre "Access to [Biological Diversity] [Genetic material]" se definía como objetivo del Convenio en materia de Acceso, al acceso a *both direct physical access to genetic material and access to information about the genetic material* (Note to facilitate understanding of issues contained in articles under consideration by sub-working group II -UNEP/BIO.DIV/WG.2/3/7 29 APRIL 1991).

En consecuencia, a criterio de Argentina la información y datos forman parte del concepto *recurso genético*, por lo que no es necesario realizar modificación alguna para incluir su tratamiento.

Cuestiones terminológicas

Con relación a la cuestión terminológica, Argentina sostiene que la terminología que mejor refleja el uso actual en el ámbito científico es “*información genética*”. Dado que el término “*digital*” supondría una sola forma de trasmisión de la información y podría haber otras no digitales y el término *secuencia* sólo refiere a la información contenida en la posición secuencial de nucleótidos en ADN o ARN o de aminoácidos en una proteína y no contempla otra información contenida en estas biomoléculas, tales como la estructura tridimensional, o grado de metilación u otras modificaciones que podrían ser fuente de información sobre la expresión y o regulación de la expresión de las mismas.

Australia

SUBMISSION

Australia thanks the Secretariat for the invitation to submit views and information on Digital Sequence Information on Genetic Resources, as communicated in Notification 2019-012

Ref: SCBD/NPU/DC/VN/KG/RKi/87804 of 5 February 2019.

Part One: Submission of views and information:

- a) To clarify the concept, including relevant terminology and scope, of digital sequence information on genetic resources and if and how domestic measures on access and benefit-sharing consider digital sequence information on genetic resources;

Australia notes more generally, there continues to be inconsistency in the terminology utilised for describing the ‘electronically held sequence information which represents the biological composition of genetic material’ across international fora – such as digital sequence information, genetic sequence data and *in silico*.

For the purposes of this submission, and in line with our submission to CBD Notification 2017-037, Australia defines “digital sequence information on genetic resources” as electronically held sequence information which represents the biological composition of “genetic material” as defined under the Convention.

Australia continues to consider digital sequence information on genetic resources (or any term used, including but not limited to genetic sequence data or *in silico*) and the physical genetic resources and material as distinct entities. Australia also considers digital sequence information on genetic resources (or any other such terminology) and ‘derivatives’ as defined under Article 2 of the Protocol as distinct entities.

This distinction is clearly articulated in the current definitions under both the Convention and the Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of the Benefits Arising from their Utilisation (Nagoya Protocol). Under the Convention and the Nagoya Protocol, ‘genetic resources’ are defined as ‘genetic material of actual or potential value’. ‘Genetic material’ is defined as ‘any material of plant, animal, microbial or other origin containing functional units of heredity’. Neither of these definitions align with the non-physical and electronic nature of digital sequence information. This distinction also aligns with the outcome of lengthy debate in the establishment of the Nagoya Protocol.

To consider digital sequence information a genetic resource under the Convention and the Nagoya Protocol would require a renegotiation of both the Convention and the Nagoya Protocol to redefine ‘genetic material’. This is because ‘information’ does not contain ‘functional units of heredity’ or genes.

Do domestic measures on access and benefit-sharing consider DSI?

The Australian Government has domestic measures in place that provide for the access and benefit-sharing of biological resources and the usage of Traditional Knowledge. These measures do not consider or provide coverage of digital sequence information.

Australia’s access and benefit-sharing legislation is guided by the *Environment Protection and Biodiversity Conservation Act 1999* (EPBC Act) and the *Environment Protection and Biodiversity Conservation Regulations 2000* (EPBC Regulations).

The EPBC Act defines biological resources and genetic resources as having physical attributes:

Biological resources: includes genetic resources, organisms, parts of organisms, population and any other biotic component of an ecosystem with actual or potential use of value for humanity.

Genetic resources: means any material of plant, animal, microbial or other origin that contains functional units of heredity and that has actual or potential value for humanity.

In line with the EPBC Act and Regulations, ownership rights to, and access and benefit-sharing requirements of biological resources depend on whether they are found in Commonwealth, State or Territory government lands or waters, indigenous lands (of which there are different types with different associated rights), freehold or leasehold lands.

For biological or genetic resources to be accessed/sampled/utilised from Commonwealth lands or waters, those seeking access must apply to the Australian Department of the Environment and Energy for a permit. Permits for access to biological resources are available for either commercial, potentially commercial or non-commercial purposes. The regulations also require the prior informed consent of the indigenous owner or native title holder, where access is to genetic resources on indigenous peoples' land and the provision of a benefit-sharing agreement must provide for reasonable benefit-sharing arrangements, including protection for and valuing of any indigenous peoples' knowledge to be used.

b) On benefit-sharing arrangements from commercial and non-commercial use of digital sequence information on genetic resources.

Australia does not support the inclusion of digital sequence information on genetic resources (or any term used, including but not limited to genetic sequence data or *in silico*) in access and benefit sharing arrangements. Australia's position is that the current open access to digital data remains in place. The generation and open sharing of digital sequence information on genetic resources provides benefits through increased scientific information and discovery. The use of digital sequence information on genetic resources increases the value of biological diversity and enables scientific progress and innovation.

Part two: Submit views and information on capacity-building needs regarding the access, use, generation and analysis of digital sequence information on genetic resources, in particular for the three objectives of the Convention.

Australia has no submission on this component of the request.

NOTE: Information provided in this response has been drawn from Australian Government agency input.

Belarus

SUBMISSION

Conclusions of the National Analytical Report “Synthetic Biology and Digital Sequence Information”

1. Organisms developed by synthetic biology techniques are currently defined under the international legislation as living modified organisms (LMOs). Such organisms fall under the scope of the Convention on Biological Diversity and the Cartagena Protocol on Biosafety to the Convention on Biological Diversity. Currently, the regulation of such LMO circulation in the Republic of Belarus and their transboundary movement should be carried out in accordance with the Cartagena Protocol on Biosafety to the Convention on Biological Diversity and the national legislation harmonized with the Cartagena Protocol. It should also be noted that such organisms are valuable genetic resources and DSI information should be considered when addressing the issues of digital information related to genetic resources.

2. Despite the fact that DSI is widely published in specialized databases or data banks, there are several practices for accessing digital information on genetic resources' sequences and such information use:

2.1. Unhindered access to information in global, regional or national databases;

2.2. Databases that require registration by their users, data access agreements, or are available only to scientific project performers;

2.3. It should also be taken into account that a number of data, such as data on molecular marker sequences, detection and identification methods, molecular genetic properties of new plant varieties, strains of microorganisms or animal breeds may be subject to patenting, and therefore not all data may be available.

3. Obtaining DSI on genetic resources requires in most cases access to genetic resources, although in some cases it may be difficult to associate digital sequence information with a genetic resource. Where a number of Parties are involved, the multilateral access and benefit-sharing process, including the cases where developers are several country Parties to the Convention on Biological Diversity and the Nagoya Protocol, can be very complex.

4. Whereas there is no consensus among the Parties to the Convention on Biological Diversity and the Nagoya Protocol on Access to Genetic Resources and Benefit-Sharing on the issue related to benefit-sharing arising from the use of DSI with regard to genetic resources, the Republic of Belarus as a Party to the Convention and the Protocol should actively participate in further discussions among countries to develop a definition for “digital information on genetic resources’ sequences”, elaborate universal regulatory mechanisms for such resources and their use on a fair and equitable basis and derive benefits arising from the genetic resources of the Republic of Belarus.

5. In doing so, the following factors should be taken into account:

5.1. There is a well-established international system for providing information on DSI and of free access over the Internet for all. Research data is currently provided by researchers globally to specialized banks or databases, and in many cases advance deposition to public databases is a requisite for publishing a scientific article. Such a system for providing information has been designed to create a barrier-free environment for access to scientific information and research, biodiversity related fundamental research in particular, and can be considered as capacity-building for country Parties to the Convention and its Protocols. Restricted access to such information, licensing bans and fees, including information on nucleotide and amino acid sequences, can lead to a decrease in the level of scientific research or duplication of such research. In addition, there are large social and community benefits arising from using and accessing DSI and this underlines the importance of publicly available databases.

5.2. At the same time, there is information on DSI, which may be subject to patenting and an exclusive right to an invention is protected by the state and proved by a patent. For example, molecular marker sequence data. The rules of patenting in the Republic of Belarus and the rules of international patenting are given in Annex 1 and can be considered in the context of DSI issues’ discussion.

5.3. The problem of kind and in-kind benefit-sharing is determined by the fact that there may be no marginal value for the use of benefits and benefit-sharing obligations may perpetuate.

5.4. Monitoring, accessing and using DSI with the involvement of several countries, holders of a genetic resource, or several countries, developers of a genetic resource, can be very difficult.

6. Commercialization of products arising from the DSI use can be considered as a triggering event for benefit-sharing and initiation of monitoring in the Republic of Belarus. Perhaps, the establishment of contracts between the executing Parties of state and international scientific projects should be considered, including the way of distributing benefits in case of developments’ commercialization. In the Republic of Belarus, at present, when executing state scientific research programs, clear plans are developed in advance to integrate the developments and they also define each organization responsible for the integration with several scientific projects’ executors. However, it may be necessary to consider signing the contract between the executing Parties of scientific projects establishing the holders of genetic resources developed during the project, including how the new genetic resource-related benefits will be distributed, and whether the research results will be patented.

Brazil**SUBMISSION**

One of the 17 megadiverse countries (SHI, SINGH, KANT, ZHU, & WALLER, 2005), Brazil is often considered the most biologically diverse country in the world. Brazil's biodiversity is an essential resource for its people, not only directly because of the environmental services it provides, but also due to the development opportunities that these represent. Since the past decade, Brazil made impressive progress in fighting threats to biodiversity in various fronts, particularly in establishing protected areas, in the fight against deforestation and in regulating sectors that either threaten the country's biodiversity endowment or propose to use it in a sustainable way across the country's various landscapes.

So far, 117,289 species of animals are known to Brazil, their clear majority being arthropods (about 85%, almost 94,000 species) and chordates representing 10% of fauna species (Ministério da Ciência, Tecnologia e Inovações e Comunicações e Ministério do Meio Ambiente, 2018).

At this moment, 46,506 species are recognized for the Brazilian flora: 4,754 of Algae, 33,109 of Angiosperms, 1,564 of Bryophytes, 5,718 of Fungi, 30 of Gymnosperms and 1,331 of ferns and Lyophytes (Jardim Botânico do Rio de Janeiro, 2018).

The Convention on Biological Diversity - CBD explicitly recognized the authority of States to determine access to genetic resources as part of their sovereign rights over natural resources under their jurisdiction. Furthermore, it obliges all contracting parties to take legislative, administrative or policy measures, to share in a fair and equitable way the results of research and development and the benefits arising from the commercial and other utilization of genetic resources.

For almost 20 years now, Brazil has put in place an ABS System and albeit the first national legislation (Provisional Act nº 2.186-16), from 2000, was revoked by the new Biodiversity Bill, (Law nº 13,123, enacted in 2015 and regulated by Decree nº 8,772, established in 2016 and in force since November 2017), the National ABS Competent Authority, the Genetic Heritage Management Council (CGEN), was maintained in charge of the coordination, development and implementation of Policies regarding access to genetic heritage (genetic resources - GR) and associated traditional knowledge - ATK and benefit sharing.

(a) To clarify the concept, including relevant terminology and scope, of digital sequence information on genetic resources and if and how domestic measures on access and benefit-sharing consider digital sequence information on genetic resources;

Since the year 2000 Brazil has an established legislation on access to genetic resources, traditional knowledge and benefit-sharing in accordance with the Convention on Biological Diversity.

The Provisional Act No 2,186-16/2001 regulated articles 1, 8, 15 and 16 of the Convention and established a governance system for access to genetic resources and to associated traditional knowledge. It created the Genetic Heritage Management Council – CGen - and devised an administrative procedure for obtaining access authorization, prior informed consent and mutually agreed terms.

The Provisional Act interpreted and matched the expression used in the Convention on Biological Diversity – CBD - “genetic resources” as the term “genetic heritage”, used in Brazil's 1988 National Constitution. Genetic Heritage was defined by the Provisional Act as the **information of genetic origin**, contained in samples of all or part of a plant, fungal, microbial or animal species, in the form of molecules and substances originating in the metabolism of these living beings, and in extracts obtained from in situ conditions, including domesticated, or kept in ex situ collections, if collected from in situ conditions, within the Brazilian territory, on the continental shelf or in the exclusive economic zone.

Thus, Brazil regulates the use of genetic information, even if disengaged from the physical sample since its first legal framework on ABS.

During the 15 years that the Provisional Act was in force, CGen granted over 2600 access authorizations and established 295 Benefit-sharing agreements (Ministério do Meio Ambiente, 2018). Including in those are some cases that have used exclusively digital sequence information.

This understanding has been maintained by the new legal framework on access and benefit sharing: Law No 13,123/2015 (Presidência da República, 2015) defines genetic heritage as the genetic information from plants, animals, and microbial species, or any other species, including substances originating from the metabolism of these living organisms.

Decree No. 8,772/2016, that regulates the mentioned Law, also required the identification of the genetic resources and its origin, including a geo-referenced coordinate of the location where the physical sample was collected *in situ*, even if obtained from *ex situ* or *in silico* sources. Therefore, Law 13123/2015 and its decree already include in its scope the use of digital genetic information, and users are subject to the need for registration and sharing of benefits from economic exploitation of products or reproductive material arising from it.

Moreover, in what relates to need for clarification on the “concept, including relevant terminology and scope, of digital sequence information on genetic resources”, it must be stressed that a systemic reading of the CBD and The Plant Treaty strongly influenced the elaboration of Law 13,123/2015 and its Decree No. 8,772/2016. The CBD defines "genetic material" as any material of plant, animal, microbial or other origin containing functional units of heredity. Moreover, according to the Oxford Dictionary, the word "material" can be defined as "information or ideas for use in creating a book or other work". On the other hand, the definition of the word "matter" is "physical substance in general, as distinct from mind and spirit; (In physics) that occupies space and possesses rest mass, especially as distinct from energy". The term "material" should not be confused with the term "matter". The definition of the word "material" allows the interpretation of the term to include the set of information associated with the genetic resource, that is, the substrate information or working material.

Therefore, it is not only conceivable to understand the word "material" in the broader scope of its meaning, but it offers a more flexible and proper meaning. To restrict its understanding to match the meaning of the word “material” to the meaning of the word “matter” is jeopardize the obligation to share benefits, the sovereignty of the countries parties over their genetic resources, the Convention on Biological Diversity and the Plant Treaty.

Even if genetic information obtained digitally is to be considered as excluded from the concept of genetic material, a systemic interpretation of the Convention on Biological Diversity and the Nagoya Protocol leaves no doubt that the use of this information is subject to benefit sharing. The means of transmission of genetic information, whether in the form of matter from a DNA sample or as information stored *in silico*, is irrelevant to the fulfillment of this obligation. Since there was a "utilization" of a physical sample to access this type of information, its application and subsequent commercialization should be shared in a fair and equitable way, in line with Article 5 of the Nagoya Protocol and article 10 of the Plant Treaty.

Therefore, the discussion of digital sequence information within the scope of international agreements ultimately does not impact the effective application of the CBD nor the Nagoya Protocol.

Furthermore, under the Pandemic Influenza Preparedness Framework, which one of its main objectives are access to vaccines and sharing of other benefits, there is already a clear definition of “Genetic sequences” which “means the order of nucleotides found in a molecule of DNA or RNA. They contain the **genetic information** that determines the biological characteristics of an organism or a virus.” In other chapter the

PIP framework establishes the procedure for best practices relating to genetic sequence data (World Health Organization, 2011).

In this context, it's easy to perceive that the object of discussion within the context of “digital sequence information” is not the word “digital”, which corresponds only to the medium in which information is transferred, and neither in the word “sequence” since it only signifies the order in which nucleotides are presented, but in its main core: the genetic information transmitted through digital media or any other media in a sequenced form or any other form. Thus, international fora discussing DSI or any other terminologies, such as "genetic sequence data", "dematerialized genetic resources", "*in silico* utilization", and "natural information", should converge in adopting “genetic information on genetic resources” as the proper terminology.

(b) On benefit-sharing arrangements from commercial and non-commercial use of digital sequence information on genetic resources.

The new ABS Legislation has been in place since November 2017, when the ABS electronic registration system “SisGen” started to operate. In general, there is no need for a prior authorization to start a research or development activity on genetic heritage. The prior authorization was substituted by a registry made with the SisGen, which must be concluded before some specific moments, as explained below.

The register is a declaratory, but mandatory instrument. The SisGen – The National System for Genetic Heritage and Associated Traditional Knowledge Management – <https://sisgen.gov.br> – is the electronic system maintained and operated by the Executive – Secretariat of CGEN, it is the “one stop shop” for the registration of ABS activities.

It manages the registry of access to genetic heritage or associated traditional knowledge; Notifications of finished product or reproductive material and benefit-sharing agreements. Additionally, the SisGen issues the Certificates of lawful access, that, to be granted, the access (research and development activity) registration must be carried out previously:

I – On the remittance of samples of genetic material;

II – On the application for any intellectual property rights;

III – On the commercialization of the intermediate product (by-products);

IV – On the disclosure of final or partial results in scientific or communication circles; or

V – On the notification of finished product or reproductive material developed as a result of the access.

Users are free to choose the best moment to do the registration as long it is before the abovementioned triggering events. Moreover, since there is no need for a prior registration, if a given access activity doesn't have any results, any intellectual property right applications, products or processes developed, that access activity doesn't have to be registered. The main idea is to promote and facilitate access and to only demand information when a concrete result has been achieved, moment in which the user must declare the access (research and technological development) and provide all the required information.

Besides the registry of access, the notification on finished product derived from access to Genetic Heritage is also a declaratory instrument made with the SisGen. The Notification precedes the beginning of any activities of economic exploitation of a finished product.

It is through the notification that users of the Genetic Heritage declare to comply with the requirements of the Law and indicates the preferred modality of Benefit-Sharing to meet their legal obligations. The modality is up to the User to decide and are “monetary”, through a payment to the National Fund, or “non-

monetary”, with the user directly funding a conservation project or activity, in accordance to the National Benefit Sharing Programme created by the Law n° 13.123/2015. In the non-monetary modality, a Benefit-Sharing Agreement must be signed, foreseen all the activities that the user declare to execute as benefit – sharing.

The Notification of a Finished Product equals to the celebration of mutually agreed terms, in accordance to the article 15 of the CBD, since the user agrees with the terms and conditions required by the national legislation. The Law predicts contract clauses, exemptions, sectoral treatment, and other terms and conditions that includes beneficiaries, users and providers rights; the types, duration and amounts due to pay as benefit sharing, as well as dispute resolution provisions and an appellate Council.

Also, in this sense, there is no different treatment or requirements regarding the notification of a finished product arising from research and technological development conducted on Digital Sequence Information. The required benefit-sharing arrangements include the economic exploitation of a finished product derived from a genetic heritage obtained from an in-silico source.

A finished product is defined by Law as a Product which is apt to be used by the final consumer, whether it is an individual or legal entity. Moreover, the Benefit sharing obligations applies only to a finished Product, that must be derived from access (research and technological development in the Brazilian Law), independently if it was produced in the country or abroad, and finally, the Genetic Heritage should be one of the main elements adding value to the product.

That means that in some cases there may be a finished product developed from a genetic resource that won't share benefits because the genetic resources presence in the finished product isn't “crucial to the existence of its functional features or its commercialization appeal”.

According to the Law, it doesn't matter who has conducted the access on DSI or who is commercially exploiting the finished product, it is the manufacturer of the finished product that must meet the Benefit-Sharing obligation.

After the registration and notification procedures, the user does not have to provide any further information, only if the Administrative verification procedure identifies any wrongdoing, which, according to the case, may be corrected directly in the System, or may demand more complex remedies, including the registry cancelation. It is the obligation of the user to provide all the true and correct information as demanded by the registration procedure and required in the Law and its regulations.

Since the Brazilian ABS model is declaratory, after the fulfillment and conclusion of the registration procedure, the SisGen provides a receipt of access registration, which constitutes a proper document to demonstrate that the user provided the information required, and has the following effects:

I – Allows for:

- a) The application for any intellectual property right;
- b) The commercialization of an intermediate product;
- c) The disclosure of results, final or partial, of the research or technological development in scientific or media circles; and
- d) The notification of a finished product or reproductive material developed as a result of access; and

II - Establishes the beginning of the Administrative Verification Procedure.

It must be stressed that the verification procedure is responsible for verifying irregularities, the user will not need to wait for the verification procedure to perform the activities mentioned above.

Finally, the CGEN issues the Certificate of Regularity, which declares the regularity of the activity up to the date of its issuing by the ABS National Authority and prevents the application of sanctions by the competent authorities in what regard to the activities carried out up to its date of issuing.

In the light of these considerations, it must be recalled that the conclusion of the registration of a Research activity by the user equals to the obtaining of a non-commercial access permit. Thereto, the registration of a Technological Development activity by the user, which in the Law is considered as a “systematic work on genetic heritage carried out with the objectives of developing new materials, products or devices, or improving or developing new processes, for economic exploitation”, equals to the obtaining of a commercial access permit. Finally, the conclusion of a Product Notification is the celebration of a Mutually Agreed Term between an User and the Ministry of Environment, which according to the Law, represents the Federal Government as the national genetic heritage provider in benefit-sharing negotiations.

If a basic research on a genetic heritage, even if obtained from *in silico* origin, becomes a Technological Development, the user doesn't have any further obligation besides updating his registry in the SisGen, what must be done once a year when better suit to the user.

Beyond, after one year and a half since SisGen started its operations, almost **800** Legal Persons (60% companies) and more than **25 thousand** natural persons concluded their registrations and are providing information on their research and development activities arising from genetic heritage (including from *in silico* origin) and ATK in the SisGen.

By May 2019, there were over **47 thousand** access (research and technological development) activities registered in SisGen, from which 16% (3,747) with declared commercial intention since were registered as technological development activities. Likewise, by May 28, there were **1,500** Finished Products Notifications in the SisGen.

Therefore, there are over 47 thousand access activities in conformity with the legislation, which means the same number of permits of access “granted” and 1,500 Mutually Agreed Terms celebrated under the new ABS Law. In addition, according to the Executive-Secretariat of the CGEN, by now, 400 Certificate of Regularity were requested by users, which are to be approved by the CGEN.

Comparing the former ABS Act, which issued 2,600 permits and more than 295 ABS Agreements/Contracts in 15 years, to the new Legal framework on ABS, the current model obtained over 200 times more regular activities and over 75 times more contracts than the old model in the same one period.

Once there is no different treatment for a genetic heritage obtained from an *in situ*, *ex situ* or *in silico* source, been subject to the same obligations and rights, the SisGen registry has specific fields in which the user can inform the origin of its genetic heritage from *in silico* origin. The SisGen allows the user to inform the name of the data base, the genetic heritage code of access from that data base and a link to the information provided regarding the genetic heritage and the data base from which it was obtained.

Out of the almost 47.000 registered access activities in the SisGen by now, **449** declared *in silico* origin, from which **64** declared commercial intention activities, through the registration of Technological Development activities arising from the utilization of digital sequence information/genetic information on Genetic Resources.

Ergo, out of those 449 registrations, 385 are equivalent to “**benefit-sharing arrangements from non-commercial use of digital sequence information on genetic resources**” as much as those 64 Technological Development registries are “**benefit-sharing arrangements from commercial use of digital sequence information on genetic resources**”.

It must be recalled that, in light with what states the Law n° 13,123/2015, an access with commercial intention (Technological Development) or non commercial intention (research) doesn't have to share benefits, although they are subject to the terms and conditions of the Law, which foresees benefit sharing arrangements and exemptions, with which the user and the provider mutually agreed upon, once the user concludes the access activity registration in the SisGen.

In order to provide concrete examples on the “**benefit-sharing arrangements from commercial use of digital sequence information on genetic resources**”, one could refer to a Technological Development activity registered in the SisGen, which proposes the use of informatics techniques to find pharmacological receptors (proteins), deposited in the Protein Data Bank PDB, of natural products from the Brazilian Biodiversity. It also declares the databases of molecules with biological activity ZINC and SEA as source of the *in-silico* origin.

Another development activity registered in the SisGen informing the data bank from which the DSI/Genetic information on genetic resource was obtained is one which aims to develop a prototype kit for confirmatory diagnosis as a complementary technique for the detection and identification of a given parasite. This technological development declared a Public Funded Antibody Platform as source of the genetic heritage.

Moreover, a third example out of 64 cases of DSI utilization is an invention that comprises the use of peptides synthetic derivatives of the toxins of a Brazilian Invertebrate for the treatment of ocular diseases. The source of the DSI utilized was “<http://www.uniprot.org>”. The UniProt Knowledgebase (UniProtKB) is “the central hub for the collection of functional information on proteins, with accurate, consistent and rich annotation”. UniProtKB mission is “to provide the scientific community with a comprehensive, high-quality and freely accessible resource of protein sequence and functional information”.

On a quick search in the UniProtKB website with the term “Brazil”, it showed almost 225,000 entries. Each entry, according to UniProtKB, captures the core data on a specific DSI, “mainly amino acid sequence, protein name or description, taxonomic data and citation information), as much annotation information as possible is added”, offering freely accessible resource of Digital Sequence Information.

Nevertheless, since there is in general no prior authorization to use Genetic Heritage from Brazil, anyone using that proteins sequences would have to register they results or notify products only when there is a concrete result and before some triggering events, such as the publication of a Scientific Paper, a Patent Application, a By Product commercialization, a finished Product Notification etc. That is why is paramount that databases require and provide the origin of the genetic information that they store and offer.

In other words, Brazilian genetic heritage can be freely accessed, but the results and products of its utilization must be regularized by a registration or notification procedure, in the proper moment and according to each case. Its paramount for Brazil to foster research and development arising from its genetic diversity and, having in mind the evolution of the techniques available to do so, it is the national understanding that access, including through the utilization of genetic resources from an *in-silico* origin, must be facilitated to generate the benefits that will fund biodiversity conservation and sustainable use. But to do so, the regulation must focus on results other than procedures.

Finally, in what relates to the verification of those registries and the information provided by the Users, done by the Administrative Verification Procedure, the total number of arrangements with possible use of digital sequence information on technological developments, as well as the information regarding the finished products developed by the utilization of DSI, may change.

Since the verification procedure review the registries to identify wrongdoing and mistakes, such as the utilization of Genetic Heritage that does not belong to the Brazilian sovereignty, some activities declared will be reclassified or corrected, with possible inclusion and withdraw of registries considered as

arrangements with possible use of digital sequence information on technological developments instead of use on non-commercial research.

Albeit the numbers may change, it is expected to raise, and, in short time, more information, including on Benefit sharing arising from the economic exploitation of finished products derived from DSI utilization, will be shared through the ABS Clearing house.

Considering the foregoing and notwithstanding the fact that the DSI subject may still cause some confusion, due to its unclear scope and different point of views, some solutions of the ABS Law No. 13,123 of 2015 could be considered for the implementation of the use of genetic information in the CBD and its Protocols, as possible ways to deal with DSI challenges.

In this sense, Brazil has adopted:

- A facilitated mechanism for access to genetic resources, with a change in the focus of regulation, previously focused on the control of access to genetic resources, now shifted towards control of the economic exploitation of products or reproductive materials arising from access;
- The development of an online registration system to trace, track and oversee access to genetic resources and associated traditional knowledge activities. The SisGen electronic system is declaratory, as opposed to the old *modus operandi* of the Provisional Act in which a procedure for validation of documents was in place;
- The registration must only be carried out prior to specific moments such as shipment, request for intellectual property rights, publication of results and commercialization. Research and development activities that do not result in any of the above-mentioned activities are not demanded to register;
- Registration is not needed prior to access (research and development) itself when only the genetic resources are accessed, without access to traditional knowledge: these activities are not restrained by any prior administrative procedure for granting access;
- Prior informed consent for access to genetic resources was granted by the National Congress: there is no administrative procedure for access to genetic resources; Prior Informed Consent for access to TK is mandatory and should be obtained directly with the ILCs;
- Economic exploitation of a finished product or reproductive material was established as the single point of incidence of benefit-sharing obligations: this is the link of the value chain with the highest value added, discharging any research and development activity: economic benefits are to be shared when they do exist;
- Because of the single point of incidence, the economic exploitation of any intermediate product is exempt from benefit sharing obligations;
- The percentage of monetary benefit sharing from products or reproductive material derived from the use of genetic resources is established as 1% of net revenues from the product or reproductive materialsales: there is no speculation of values and no surprises for genetic resources users. It gives predictability and legal certainty to invest in Biobased products arising from access;
- The clearly established point of incidence combined with a defined percentage of benefit-sharing to be valued under a specific concept such as “net revenue” make the monitoring of compliance feasible, since they are based on fiscal and accounting principles and rules;
- When the user chooses to share the benefits through non-monetary means, such as a conservation or social project, benefit-sharing is equivalent to seventy-five percent of the predicted value for the monetary modality. This concession considers expenses the user might have in implementing the project and encourages the non-monetary modality;
- Licensing, transferring or permitting any use of intellectual property rights does not require benefit sharing. Benefit sharing obligations exist only when a finished product or reproductive material using the licensed intellectual property is commercialized to the final consumer;
- Micro-businesses, small businesses, micro individual entrepreneurs, traditional farmers and their cooperatives are exempt from benefit-sharing obligations;

- Another solution was the establishment of a National Benefit-Sharing Fund for centralization and subsequent redistribution of benefits arising from the use of genetic resources and associated traditional knowledge through a management committee for actions focused on research, development and conservation of genetic resources and protection of associated traditional knowledge;
- Once the due amount to be shared is given by law (1% of net revenues from the product or reproductive material sales), users can pay the benefits directly to the Fund, through an electronic voucher provided by SisGen, once the registration and notification requirements are fulfilled and when there are benefits to be shared. The need to sign the Benefit-Sharing Agreement (MAT) will occur only when users decides for the nonmonetary modality.

Furthermore, many have pointed the difficulties in identification of both the genetic resource or its origin as an argument for preventing digital sequence information from being considered within the scope of both the CBD and the Plant Treaty. Brazilian Decree 8,772/2016 has already foreseen procedures to be adopted in cases of techniques which access microorganisms that are not isolated from a specific substrate and have not been identified, such as metagenomics.

Brazil also have positioned in favor of using the Global Multilateral Benefit-Sharing Mechanism to resolve issues of benefit sharing relating to situations in which prior informed consent cannot be obtained, such as lack of origin information, transboundary situations or products and reproductive material resulting from multiple access from different origins (Ministry of Foreign Affairs of Brazil, 2017).

Therefore, useful instruments are already in place to resolve issues for the use of digital sequence information within the framework of the CBD. There are viable regulating strategies and establishment of trigger points that will not impede the rapid share of information, crucial for our current scientific demands.

Final remarks

The use of genetic information in the context of access and benefit sharing is regulated by the Brazilian legislation since 2000.

In line with the definitions of the Brazilian Law, the object of international discussions should not focus on digital sequence data but in its core object: the genetic information contained by it.

There are several alternatives to regulate access to genetic information, including for food and agriculture, that is not by creating overly bureaucratic Sisyphean tasks. Mainly, countries should change the focus from regulating of processes/procedures towards regulating results. This shift relieves the bureaucratic burden of research and development and focuses on the end of the chain, the economic exploitation of products and reproductive material.

When national legislations focus on monitoring end-products, instead of monitoring the process to obtain those, more user-friendly ABS systems will come to exist, strengthening confidence in ABS international system.

For that countries should invest in the creation of simple, declaratory and transparent regulations, but at the same time invest in effective tracking and tracing tools that allows monitoring of compliance. Additionally, it should provide changing of intention mechanisms (specially from non-agricultural research to agricultural research, and vice-versa), a clearly defined triggering event, quantifiable and non-speculative values for benefit-sharing, preferably based on fiscal and accounting principles and rules, and a strategy in which monetary benefit-sharing should be an obligation when there is a clear monetary benefit being obtained from the use of a genetic resource.

Predictable rules will allow users to foresee their costs and obligations, in the short and long term, and will provide legal clarity to users and thus encourage the use of genetic resources. Legal measures that facilitate

and foster research and development will generate more benefits, which can be channeled to biodiversity conservation and sustainable use, fulfilling the objectives of the international agreements on ABS.

Lastly, Brazil firmly believes the concept of ABS is a useful and workable tool and that it represents a potentially substantive source of national funding for conservation in many countries. ABS is a powerful regulatory tool to promote synergies, cooperation, business and social development among countries, and societies moving towards a different sustainable, efficient and democratic model of economic development. It encodes a new understanding on how to do business, guided by environmentally-sound practices, respect and social responsibility, what allows the rational economic exploitation of biodiversity to finance its conservation and sustainable use.

References

Jardim Botânico do Rio de Janeiro. (2018). Retrieved from Flora do Brasil 2020 em construção:
<http://floradobrasil.jbrj.gov.br>

Ministério da Ciência, Tecnologia e Inovações e Comunicações e Ministério do Meio Ambiente. (2018). *Sistema de Informação da Biodiversidade Brasileira*. Retrieved from Catálogo Taxonômico da Fauna do Brasil: <http://fauna.jbrj.gov.br>

Ministério do Meio Ambiente. (2018). Retrieved from Atividades do CGen durante a vigência da MP nº 2.186-16/2001: <http://www.mma.gov.br/patrimonio-genetico/conselho-degestao-do-patrimonio-genetico/atividades-do-cgen-durante-a-vigencia-da-mp-n-2-186-16-2001/relatorio-de-atividades>

Ministério do Meio Ambiente. (2018). Retrieved from Biodiversidade:
<http://www.mma.gov.br/biodiversidade>

Ministry of Foreign Affairs of Brazil. (2017). Retrieved from <https://www.cbd.int/abs/DSIviews/Brazil-DSI.pdf>

Oxford University Press. (2018). *Oxford Dictionary of English*. Retrieved from
<http://www.oxforddictionaries.com>.

Presidência da República. (2015). *Brazilian Biodiversity Law*. Retrieved from
<https://absch.cbd.int/database/record/ABSCH-MSR-BR-238963>

Presidência da República. (2016). *Decree 8,772*. Retrieved from
http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/decreto/D8772.htm

SHI, H., SINGH, A., KANT, S., ZHU, Z., & WALLER, E. (2005). Integrating Habitat Status, Human Population Pressure, and Protection Status into Biodiversity Conservation Priority Setting. *Conservation Biology*, 1273–1285. doi:doi:10.1111/j.1523-1739.2005.00225.x

World Health Organization. (2011). *Pandemic influenza preparedness framework for the sharing of influenza*. Retrieved from PIP Framework:
http://apps.who.int/iris/bitstream/10665/44796/1/9789241503082_eng.pdf

Canada

SUBMISSION

(a) Views and information to clarify the concept of digital sequence information on genetic resources

Canada supports discussion on the relationship between “digital sequence information on genetic resources” (DSI) and the objectives of the Convention, within the CBD, and recognises that “DSI” is an increasingly vital tool for the implementation of the Convention, contributing to Articles 7, 8, 9, 10, 12 and 17. While DSI is uncovered/generated by research conducted on genetic resources, Canada takes a firm position in many fora (CBD, ITPGRFA, CGRFA, WHO) that DSI is not equivalent to tangible genetic resources or genetic material, and continues to resist language that implies any such equivalency.

Regarding terminology, Canada is uncomfortable with the use of the term “digital sequence information on genetic resources”, which was introduced by CBD decisions and not, until these discussions, used by the scientific community. In summary, we consider that it is ill-defined; too general/broad to be operationally meaningful; and it can have a variety of interpretations under different contexts.

It is unclear as to what “DSI” is intended to refer. For example, does it solely refer to complete biomolecule strings, or also to partial strings? Does it solely refer to hereditary biomolecule strings, or also non-hereditary strings (such as amino acid strings)? To naturally-occurring biomolecule strings, or also non-natural biomolecules (e.g. lab-modified biomolecules; hypothetical biomolecules (lab artifacts, instrument errors); recent non-biological computer data sets (binary codes) being stored as DNA for efficient digital information storage, aka “DNA digital data storage”)?

Unpacking the term “digital sequence information on genetic resources”, we note:

- i) Regarding “digital”, the use of “digital” implies that these resources are no longer in their initial biomolecular format; further, the fact that genetic sequences can be conveyed digitally is of secondary importance. The “digital” aspect only refers to a particular mode of storage or mode of transmission and such sequences listed on a sheet of paper would still be genetic sequences.
- ii) Regarding “sequence”, the sequences in question are biomolecular sequences, not (for example) mathematical sequences or sequences of events, so there needs to be a descriptor for the type of sequence.
- iii) Regarding “information”, the term is problematic as a term in this context:
 - “information” is generally achieved after data collection and analysis;
 - various degrees of “information” can be achieved depending on when/how much/the way data is processed,
 - “information” outputs can change as new analyses are performed;
 - while original data stay the same, “information” can be outputs on a continuum such as discrete packets (genes) or broader aggregates (entire metabolic pathways);
 - how (i.e. at what state of data processing) “information” is generated usually needs to be specified to be meaningful.
- iv) Regarding “genetic resources”, recognizing that this term continues to provoke debate, Canada firmly holds nonetheless that the negotiated and agreed CBD definition applies to tangible matter, and the definition’s wording “containing functional units of heredity” clearly points to the involvement of DNA and RNA rather than other types of biomolecules (such as amino acids).

We note also that although the terms “genetics” and “genomics” are sometimes used interchangeably, they have different meanings: genetics is the science of inheritance and genomics is generally considered to be study of a complex set of genes, their expression, and how they interact with other genes and their environment.

The consideration of public health approaches to this issue has potential to contribute clarity to the CBD discussions, because word meanings and terminology must be explicit for operational public health and biomedical science exchange. “Genetic sequence data” is the term employed by the WHO PIP Framework and defined in a relatively clear and practical manner as “the order of nucleotides found in a molecule of

DNA or RNA... contain[ing] the genetic information that determines the biological characteristics of an organism or a virus”.

Operationally, public health practitioners and biomedical science tend to use more granular / operational language to refer specifically to the type of “digital information”. For example:

- *Hereditary* genetic sequences can be acquired in part (‘partial’ sequences) or to completion (‘complete’ sequences). Is the sequence a subcomponent sequence (e.g. sequence for a gene? A metabolic pathway? Or a chromosome?) vs. Is the hereditary sequence for an entire organism, namely its intact genome?
- *Controlled Vocabulary* terms should also be used to fully describe the level of *quality* for each genetic sequence data resource. Operationally, this is helpful for conveying the degree of data accuracy, level of curation and the type of curation. For example, when describing a *genome* resource, there are important QUALIFYING terms applied, such as:
 - “standard draft” (describes the minimum information needed for submission to a public database); “high quality draft” (describes sequences with little to no manual review); “improved high quality draft” (describes data, which has undergone review by people or by machines to some extent to indicate that most of the genetic data is assembled correctly, but some errors may still be present); “annotation-directed improvement” (in which genetic information in various gene regions is represented as accurately as possible); “noncontiguous finished” (includes sequences that have been reviewed by both people and machines and is considered complete except for “recalcitrant regions” that are proving problematic for genome closure); “finished” (which describes un-gapped complete genome sequences that have minimal errors, if any³).

Canada thus prefers the terms “genetic sequence data” or “genomic sequence data”, both of which, from a science operations viewpoint, better describe what “is” the sequence, and the codified and transmissible nature of that sequence. There should also be qualifying/subheading statements such as “partial”, “complete” sequence, etc. We shall use this term “genetic sequence data” (GSD) in the rest of the document.

Regarding domestic ABS measures with relevance to GSD, at the federal level, Canada does not have laws or regulations specifically targeting ABS, but the Government of Canada is committed to openness, transparency and information sharing through the Open Government Partnership. Since joining the Open Government Partnership, a multilateral initiative to foster greater transparency and accountability, Canada has developed three national action plans aimed at enhanced access to information and expanded open data, among other things. The Ministry of Environment and Climate Change Canada is the federal lead on the Open Science element of Canada’s 2016-2018 Action Plan on Open Government.

(b) Views and information on benefit-sharing arrangements from commercial and non-commercial use of digital sequence information on genetic resources

As GSD is effectively obtained via the utilization (the conducting of research...) on genetic resources, it should be regarded as a result of such utilization. The sharing of these results (the GSD) via access to databases and technology should thus be regarded as a valuable form of non-monetary benefit-sharing, shared openly and globally or via benefit-sharing arrangements as set out in mutually-agreed terms to cover the tangible genetic resources.

Canada is committed to open data and information sharing, but the extent to which such open sharing is, or can be, feasibly tracked – for example to identify and/or analyze which entities are accessing GSD and are thus beneficiaries of such sharing – varies according to the purpose and resources of the databases or

³ Source: The Genomics Standards Consortium: <https://press3.mcs.anl.gov/gensc/>

institutions concerned. Data tracking/tracing mechanisms are more of a foundational consideration for public health laboratory management systems, for example, than natural history collection management systems.

A few examples follow of Canadian institutions and projects that generate and share GSD for global conservation, sustainable use and public health-related benefit:

The Barcode of Life Data System (BOLD; <https://boldsystems.org>) is an open-access cloud-based data storage and analysis platform designed to support the assembly and use of DNA barcode data, developed at and coordinated by the Centre for Biodiversity Genomics (CBG) at the University of Guelph. BOLD is heavily-used, evidenced by the 7M+ barcodes that it hosts, representing 600K+ species, and its 25,806 users in 144 countries. Although data sharing and collaboration are primary BOLD goals, the platform adopted stringent security policies to ensure the privacy of its users. Information contained in BOLD data records (including DNA barcode sequences) is not generally disclosed through BOLD unless the corresponding project or dataset is published by its contributor. Sequence data contained in BOLD is used by the BOLD Identification Engine to provide DNA-based taxonomic identifications to public users, and provenance data and images submitted to BOLD become publicly visible through the public BOLD Taxonomy Browser. BOLD regularly submits record updates to GenBank (which are mirrored on the other databases in the International Nucleotide Sequence Database Collaboration, INSDC); while records submitted remain private on GenBank for one year to allow time for submitters to publish their findings first, BOLD encourages its users to initiate pre-publication release of their data. As well as hosting BOLD, CBG serves as the Secretariat for the International Barcode of Life Consortium, a research alliance of 30+ nations that aims to deliver a DNA-based identification system for all animals, plants and fungi as part of the long-term goal of establishing an Earth observation system for biodiversity (<https://ibol.org>). As a provider of sequencing services, CBG itself uses an MTA to clarify ownership, use and storage of genetic resources and GSD it handles. Through a series of collaborative initiatives with the CBD Secretariat and Canada-Americas Trade-Related Technical Assistance Program, and in partnerships with research institutions, government agencies and other organizations worldwide, CBG has facilitated the dissemination of scientific and technical advances in DNA barcoding, partly through capacity-building events in developing countries and training courses at CBG.

Canada is a voting participant of the Global Biodiversity Information Facility (GBIF). The Canadian Museum of Nature (CMN), a part of the Governing Board of GBIF since its earliest days, is a strong supporter of freely distributed biodiversity data. Over 30% of its specimen information, from botanical, zoological and palaeobiological collections from Canada (75%) and other countries, is now in digital format (>800K data points) and its goal is to completely digitize key information about its specimen holdings. It regularly updates its electronic specimen data at the GBIF portal – last year over 1B CMN data points were downloaded freely through GBIF – and intends to do so with GSD related to its specimens when this kind of data is accommodated by GBIF's distributed system in future. The CMN has a DNA sequencing facility on site, purchases sequencing services for partial or complete genome sequences that are related to its collections, and contributes GSD via BOLD. The CMN's own research is non-commercial and all loans are for non-commercial scientific research or education purposes; much of the CMN's research can be applied to conservation practices. Any process that restricts current museum practices of loaning specimens and sharing data, including information about genetic sequences, will seriously disrupt the scope of specimen-based research and international collaboration.

Canada's federal government Genomics Research and Development Initiative (GRDI) supports genomics research in eight federal science departments and agencies. As well as supporting development and maintenance of Canadian agricultural and public health genomics, GRDI also supports key environmental programs such as the EcoBiomics project, co-led by Agriculture and Agri-Food Canada (AAFC) and Environment and Climate Change Canada. This project uses metagenomics approaches to develop new knowledge to improve water quality and soil health by comprehensively characterizing aquatic microbiomes, soil microbiomes and invertebrate zoobiomes in varying habitats, and testing hypotheses to

enhance environmental monitoring, assessment and remediation activities. It aims to establish comprehensive baselines for assessing future changes to water and soil biodiversity at key long-term environmental monitoring sites in Canada. It uses known GSD from sources such as GenBank and BOLD to identify the taxa in the samples and potentially their role in the ecosystem being studied, but also contributes to the identification and naming of taxa that are not known to science, and shares this knowledge openly. Such biomonitoring is essential to understand taxa and communities that may need conserving and thus potential regulatory action. There is also a strong economic tie to the production systems involved, as maintaining diversity is typically correlated to healthy systems and thus production, as well as monitoring for alien invasive species and functional changes that could damage productivity. The EcoBiomics project both benefits from and contributes to global GSD-sharing. Without GSD sharing, reference sequences will be internalized and there will be essentially no libraries available to carry out identification and quality control. This would obstruct biodiversity research at national and international levels, especially on microbial and invertebrate life, and have a ripple effect to industry in Canada and elsewhere, and the production systems they serve. Such projects would no longer be possible if access to reference libraries of GSD and associated annotations becomes complicated and expensive.

AAFC also conducts research to document pollinator diversity in our agricultural landscapes and studying the impacts of land use, pesticides and pathogens on bee populations. This work involves sampling large numbers of bees, which must all be identified to species level. For such identification, AAFC often sequences various genomic regions of these bees and compares the sequences to those that have been made available in online databases through collaborative efforts of the scientific community. Access to GSD of bees from other countries is vital to be able to detect any newly introduced species; researchers also make use of worldwide GSD of pathogens known to negatively impact bees in other parts of the world, in order to screen bees sampled in Canada. The GSD from the thousands of bees and pathogens that we have produced and shared are invaluable tools for other researchers worldwide.

GSD allows for rapid distinguishing of species that are potentially invasive and harmful from those that are beneficial and part of natural ecosystems. Pesticides are often the primary short-term solution for crop protection against invasive insects. However, alternative pest management solutions are urgently needed as public concern over pesticide use continues to rise, and as many pesticide chemistries are phased out in Canada. Science-Based Departments and Agencies funded by GRDI, as well as partners such as CABI, conduct research to identify organisms from all life stages through the development of new genomics tools for efficient detection of quarantine and invasive species, thus substantially bolstering Canada's operational strategy for both prevention and effective eradication or mitigation of new invaders. Through the Quarantine and Invasive Species Project (2016-2011), 27 standard operating procedures were developed for DNA extraction for DNA barcoding, partial genome and whole genome sequencing of specific quarantine and invasive species and were transferred to the Canadian Food Inspection Agency for diagnostic purposes. This project is central for federal agencies working on invasive alien species, as it provides a diagnostic tool for the early detection (significantly lowering risk of establishment), surveillance and management of thousands of species. GSD and other information is shared openly and publicly via BOLD, GenBank and GBIF, and AAFC is developing a biodiversity portal to facilitate the open sharing of these reference sequences.

The identification of plant specimens (and, specifically, of wood and forest products) in international trade depends on access to GSD and the effective international sharing of such information. This includes the identification of regulated species, of regulated geographic populations of species, and of individual specimens where legal acquisition is in doubt. Such information also provides forensic evidence to support enforcement of international agreements that regulate trade in wildlife specimens, such as CITES. Currently, Natural Resources Canada – Canadian Forestry Service researchers are using genomic marker data from well-characterized forest tree species to develop timber identification methods that will contribute to international efforts, including by the Global Timber Tracking Network (GTTN; globaltimbertrackingnetwork.org) to address illegal harvesting and trade of forest products. The voluntary exchange of GSD (and of vouchered samples required to verify taxonomic accuracy) through international

cooperation between scientific institutions (herbariums, xylaria, etc.) is, therefore, a constructive tool that can be used to support and preserve biodiversity in countries of origin.

Present day biomedical and public health bodies are actively discussing how to make GSD more openly available to benefit humankind. This is important because nucleotide sequences contain crucial information about the structure and specific properties of an organism, including virulence or transmissibility characteristics. Whole genome sequencing (WGS) is modernizing foodborne disease surveillance, providing more accurate, higher resolution information than earlier methods. Since 2017, Canada has been generating and using GSD for routine surveillance, outbreak detection and outbreak response via the PulseNet Canada network, a data-sharing collaboration of Canadian public health and food safety agencies and labs. GSD is generated by the National Microbiology Laboratory (NML), the Canadian Food Inspection Agency (CFIA), and by Provincial Public Health Laboratories, and routinely shared among all Federal, Provincial, and Territorial partners involved in foodborne disease surveillance and outbreak response (laboratory and epidemiology), including Health Canada, and the CFIA. Data sharing has been supported and flowing in this way since 2004, when Pulsed-Field Gel Electrophoresis (PFGE) was used rather than GSD. As of 2019, GSD generated from foodborne disease surveillance in Canada will be made publicly available (with proper protections of patient privacy) on the INSDC. This is a result of several years of collaborative work with NML and all provincial public health laboratories to develop best practices and procedures to make GSD publically available, greatly increasing the value added of these data for broader public health and food safety study. PulseNet Canada is part of the larger PulseNet International community, which works together to share data, standardize surveillance tools, and build capacity in 88 countries⁴.

Sharing GSD has provided benefits to the wider research community leading to more expertise to solve public health problems. For example, through a GRDI-funded collaboration between federal and academic researchers, the Food and Water Safety project aimed to capture the genomic diversity of pathogenic E. coli from human, food, animal and environmental sources in Canada from 1980 to 2013. The generated GSD was made publically available through the INSDC and provides a huge resource of the full One Health spectrum of disease transmission. Generating and sharing GSD was a critical component of the 2010 Haiti Cholera outbreak investigation in partnership with the US Center for Disease Control (CDC). This was a high profile international instance where the power of WGS was used to inform the source of the outbreak. WGS data from this investigation was shared with the CDC and other international collaborators and deposited into INSDC.

In conclusion, Canada views GSD-sharing as an important form of non-monetary benefit-sharing. The various applications of globally-shared GSD, such as those included in this submission, make increasingly critical contributions to the implementation of the Convention at national and international levels, as well as other international agreements, bringing benefits to national and global society. While providers and users may negotiate mutually-agreed terms for the utilisation of genetic resources accessed according to domestic ABS measures, and these terms may set out how benefits, including GSD, resulting from such utilisation may be used and shared, Canada strongly supports the free exchange of GSD. It is Canada's view that rather than restricting access to GSD, emphasis and effort should be expended on bioinformatics collaboration and capacity-building to better enable all countries to use globally-shared GSD in support of their national conservation and sustainable use priorities and obligations.

Colombia

SUBMISSION

⁴ PulseNet International: Vision for the implementation of whole genome sequencing (WGS) for global food-borne disease surveillance: <https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2017.22.23.30544>.

En atención a la solicitud de la Secretaría del CDB en relación a la notificación 2019-012, asociada a la decisión adoptada en la COP14 sobre Información digital de secuencias sobre recursos genéticos, damos respuesta en los siguientes términos:

- a) Aclarar el concepto, incluida la terminología y el alcance relevantes, de la información de secuencia digital sobre recursos genéticos y si las medidas nacionales sobre acceso y participación en los beneficios consideran la información de secuencia digital sobre recursos genéticos;

Concepto de secuencia digital de recursos genéticos: son aquellas secuencias que se han generado a partir de una secuencia genética homóloga a un recurso genético *in vivo* por medio de herramientas moleculares, biotecnológicas y computacionales.

Las secuencias genéticas digitales obtenidas a partir de un recurso genético (secuencias de ADN y ARN) *in vivo*, por medio de tecnologías de secuenciación y bioinformáticas, son exactamente homólogas. Por lo tanto, la información genética contenida en el recurso genético es la misma que se encuentra en la secuencia genética digital, es por esto que para la utilización de los recursos genéticos *in vivo* o los recursos genéticos digitales deben ser aplicables las disposiciones y objetivos del Convenio de Diversidad Biológica y el Protocolo de Nagoya sin ninguna distinción.

Actualmente la legislación colombiana en materia de acceso, considera las secuencias genéticas digitales como recursos genéticos y por lo tanto dentro de los contratos de acceso a recursos genéticos y sus productos derivados se incorpora una cláusula que establece que en caso de pretender entregar a cualquier título, publicar o liberar información genética y/o química con potencial de bioprospección, aprovechamiento industrial y aprovechamiento comercial, obtenida del acceso a los recursos genéticos y sus productos derivados, se deberá previamente solicitar la autorización al Ministerio de Ambiente y Desarrollo Sostenible.

Además, en caso de pretender utilizar una secuencia genética digital derivada de un recurso genético *in vivo* de una especie nativa de Colombia y que se encuentra en una base de datos de carácter público o privado, se deberá contar con el contrato de acceso a recursos genéticos y sus productos derivados y realizar la respectiva distribución de beneficios.

- b) sobre acuerdos de participación en los beneficios derivados del uso comercial y no comercial de información de secuencia digital sobre recursos genéticos.

En la actualidad hemos firmado contratos de acceso a recursos genéticos y sus productos derivados, en los cuales se permite la liberación de secuencias genéticas digitales en bases de datos públicas y privadas siempre y cuando se divulgue el origen colombiano del recurso genético, hasta ahora no hemos tenido un contrato de acceso en el cual se pretenda utilizar una secuencia digital para uso comercial o no comercial. Sin embargo, el mecanismo para establecer el consentimiento informado previo, los términos mutuamente acordados y la distribución de beneficios para la utilización de secuencias digitales se encuentra establecida.

De conformidad con la decisión 14/20, párrafo 10 y en atención a la invitación que hace la Secretaría de presentar información sobre las necesidades de creación de capacidad en relación con el acceso, uso, generación y análisis de información de secuencia digital sobre genética. recursos, en particular para los tres objetivos del Convenio; se considera fortalecer el establecimiento de cláusulas para la autorización del acceso a las secuencias digitales y los mecanismos para la distribución de beneficios, teniendo en cuenta que los desarrollos normativos y técnicos están diseñados para el acceso a los recursos genéticos tal y como se establece en las disposiciones y objetivos del Convenio de Diversidad Biológica.

Por último, es necesario evaluar como la propiedad intelectual se relaciona con el uso de secuencias digitales en el marco del acceso a los recursos genéticos y la distribución de beneficios. Los países parte del Convenio de Diversidad Biológica deberían tener proyectos regionales de fortalecimiento de capacidades en temas como secuencia digital de recursos genéticos, biología sintética, distribución de

beneficios y propiedad intelectual que están relacionados con la administración del recurso genético y las declaraciones de la agenda post 2020.

Costa Rica

SUBMISSION

Al respecto, el Gobierno de Costa Rica considera pertinente presentar las siguientes opiniones e información:

(A) Para aclarar el concepto, incluyendo terminología relevante y alcance, de la información digital de secuencias de recursos genéticos y si y cómo las medidas domésticas en materia de acceso y distribución de beneficios consideran la información de secuencia digital en recursos genéticos;

Respecto al concepto, Costa Rica sigue la misma línea de pensamiento de la conclusión del Grupo Especial de Expertos Técnicos (AHTEG) en Materia de Información Digital de Secuencias de Recursos Genéticos (DSI, por sus siglas en inglés), sobre lo que se debe incluir dentro del concepto de DSI, lo cual se puede incluir en dos categorías:

1. Información que indique de alguna manera la composición genética o bioquímica de los recursos genéticos
 - a) Las lecturas de secuencias de ácido nucleico y datos asociados;
 - b) La información sobre el montaje de secuencias, su anotación y la cartografía genética. Esta información puede describir genomas enteros, genes individuales o fragmentos de estos, códigos de barra, genomas de organelas o polimorfismos de un solo nucleótido;
 - c) Información sobre expresión génica;
 - d) Datos sobre macromoléculas y metabolitos celulares;
2. Datos observacionales que aportan información contextual (metadatos)
 - e) Información sobre relaciones ecológicas y factores abióticos ambientales;
 - f) Función, por ejemplo, datos de comportamiento;
 - g) Estructura, incluidos datos morfológicos y fenotipo;
 - h) Información relacionada con taxonomía;
 - i) Modalidades de uso.

A Costa Rica le parece que se debería incluir explícitamente en la primera categoría las secuencias de aminoácidos y en la segunda categoría los metadatos asociados a la secuencia.

Para Costa Rica es una preocupación ¿cómo las Partes podrían dar trazabilidad a la información incluida en la segunda categoría?, la cual podría ponerse a disposición del público en general por medio de artículos científicos, informes y otros medios de libre acceso.

Respecto al alcance, Costa Rica considera que la información digital de secuencias está directamente vinculada a los recursos genéticos y bioquímicos y, por lo tanto, se encuentra dentro del ámbito de aplicación del Convenio y del Protocolo de Nagoya. En el artículo 5 del Protocolo de Nagoya está explícito que "... los beneficios que se deriven de la utilización de los recursos genéticos, así como las aplicaciones y comercialización subsecuente, se compartirán de manera justa y equitativa...".

Respecto a las medidas domésticas, en la legislación nacional costarricense, las propiedades genéticas y bioquímicas de los elementos de la biodiversidad son bienes de dominio público y corresponde al Estado autorizar su utilización. El análisis y uso de la información digital de secuencias es considerado por Costa Rica un tipo de utilización posterior de los recursos genéticos o bioquímicos.

Además, en la Ley de Biodiversidad de Costa Rica, #7788, se incluye el conocimiento dentro de la definición de acceso a los recursos genéticos y bioquímicos, la cual se transcribe literal de la siguiente manera: “Acción de obtener muestras de los elementos de la biodiversidad silvestre o domesticada existentes, en condiciones *ex situ* o *in situ* y **obtención del conocimiento asociado**, con fines de investigación básica, bioprospección o aprovechamiento económico.” (énfasis agregado).

En Costa Rica se regula el acceso a recursos genéticos y bioquímicos y se considera la información resultante como productos de la investigación. Para investigación básica y bioprospección (usos no comerciales) no se regula el uso de las secuencias, entendiéndose que no se va a solicitar un Consentimiento Previamente Informado ni Condiciones Mutuamente Acordadas con los proveedores de los recursos de donde se originaron esas secuencias, principalmente debido al alto número de transacciones que esto implica y la dificultad de encontrar al proveedor de cada una de las secuencias; y, por otro lado, teniendo en cuenta que la gran mayoría de los beneficios de este tipo de acceso es de índole no monetario, con beneficios para la conservación y uso sostenible de la biodiversidad. De esta manera se facilita el acceso para este tipo de investigación no comercial, sin embargo, la Autoridad Nacional Competente puede restringir la publicación de esas secuencias en casos específicos, con su respectiva justificación, esto para evitar que esta información entre en el mundo de secuencias disponibles en bases de datos de libre acceso y pueda ser utilizada por terceros con fines comerciales sin que haya una distribución justa y equitativa de beneficios con el proveedor original del recurso. Se debe hacer un protocolo de análisis para definir los casos.

Para uso comercial, Costa Rica considera que sí se debe asegurar que haya una distribución de beneficios. Para usos comerciales puede que el número de secuencias a utilizar sea mucho menor, lo que facilita la parte operativa. Por otro lado, se considera que el mecanismo mundial multilateral podría ser una opción para solucionar el problema de número de transacciones en estos casos.

Además, Costa Rica considera que el trabajo conjunto con entidades que manejan bases de datos de libre acceso es de gran importancia. Es indispensable que los metadatos asociados a las secuencias, incluya información clara del origen de la muestra de donde se obtuvo la secuencia. Asimismo, sería importante que los sistemas de las bases de datos de libre acceso incluyan el historial de descargas de las secuencias, lo que permitiría la trazabilidad de uso de dichas secuencias y que este historial pueda ser accesado al menos por las Autoridades Nacionales Competentes de las Partes.

(B) Sobre los contratos de acceso y distribución de beneficios para uso comercial y no comercial de la información digital de secuencias de recursos genéticos.

Es muy claro que el libre acceso y uso de información de secuencias digitales para investigación no comercial tiene por sí mismo distribución de beneficios no monetarios, los cuales se traducen en aportes importantes de generación de información para la conservación y uso sostenible de la biodiversidad, contribuyendo así al cumplimiento de los dos primeros objetivos del Convenio. Sin embargo, cuando se trate de usos comerciales, además de los beneficios no monetarios, se debe asegurar también la distribución de beneficios monetarios entre el usuario y el proveedor del material que dio origen a las secuencias.

Además, es conocido que el acceso a la información digital sobre secuencias que se mantiene en bases de datos públicas no está sujeto a requisitos de consentimiento fundamentado previo y que durante la investigación se utiliza una gran cantidad de accesiones. Por lo anterior, con el fin de no obstaculizar la investigación no comercial, y basados en el artículo 8 del Protocolo de Nagoya, Costa Rica está a favor de facilitar el acceso a las secuencias para investigación no comercial, recalcando que la distribución de beneficios se traduce en la generación de información útil para la conservación y uso sostenible de la biodiversidad.

Por otro lado, se reconoce la importancia de asegurar beneficios monetarios cuando se trate de uso comercial, por lo tanto, se debe buscar la forma de regular el acceso. Es indispensable que al momento de publicar las secuencias en bases de datos de uso libre se incluya, dentro de los metadatos de la secuencia,

el país de origen de la muestra de donde fue obtenida la secuencia. Para esto, los administradores de las bases de datos deben incluir un campo sobre el origen de la muestra de donde se obtuvo la secuencia. Además, al igual que Costa Rica, es importante que los países analicen los casos concretos en los que sea necesario restringir la publicación de secuencias, definiendo y siguiendo una ruta para el análisis de los casos.

Necesidades de creación de capacidad en relación con el acceso, uso, generación y análisis de información digital de secuencias de recursos genéticos, en particular para los tres objetivos del Convenio.

Capacitación en bioinformática para poder realizar minería de datos (data mining) y búsqueda de información masiva; en generación de cláusulas modelo que regulen el acceso y uso de la información digital de secuencias de recursos genéticos y bioquímicos; uso de las bases de datos de libre acceso.

Ethiopia

SUBMISSION

In a notification dated February 7, 2019, the CBD Secretariat, pursuant to decisions 14/20 and NP-3/12, respectively, invited Parties, other Governments, indigenous peoples and local communities, relevant organizations and stakeholders to submit views and information:

- (a) To clarify the concept, including relevant terminology and scope, of digital sequence information on genetic resources and if and how domestic measures on access and benefit-sharing consider digital sequence information on genetic resources;
- (b) On benefit-sharing arrangements from commercial and non-commercial use of digital sequence information on genetic resources.

In addition, Parties, other Governments and indigenous peoples and local communities were invited to submit information on “their capacity-building needs regarding the access, use, generation and analysis of digital sequence information on genetic resources, in particular for the three objectives of the Convention.

Ethiopia as a party to CBD submits the following views and information on DSI:-

- (a) Views and information to clarify the concept, including relevant terminology and scope, of digital sequence information on genetic resources and if and how domestic measures on access and benefit-sharing consider digital sequence information on genetic resources.

1. Terminology

DSI is not an appropriate terminology to capture the various types of information on genetic resources that may be relevant to the three objectives of the CBD. Therefore, we suggest using the term **genetic information or genetic sequence data**.

2. Scope

Any information that results from the utilization of genetic resources should fall under the scope of ABS legislation and subject to benefit sharing

3. how domestic measures on access and benefit-sharing consider digital sequence information on genetic resources

The revised draft Ethiopian ABS proclamation incorporates DSI in its scope and definition of genetic resources. ‘Genetic resource’ is defined as any material of biological resource containing genetic information having actual or potential values for humanity and includes derivatives and digital sequence information. PIC and MAT requirements that apply to access GRs also apply to DSI and apply also benefit sharing obligations from use of DSI for both commercial and non-commercial use. The details will elaborate in the Ethiopian ABS regulation under review. The Mutually Agreed Terms and the permit templates will be revised to incorporate mandatory clauses that address on conditions to use genetic information that results from utilization of GRs, whether it is in the public or private databases.

(b) Views and Information on benefit-sharing arrangements from commercial and non-commercial use of digital sequence information on genetic resources.

Ethiopia suggests that there should be an obligation of benefit sharing from use of DSI that result from utilization of GRs. The benefit sharing should be from commercial and non-commercial use of DSI.

(c) Information on capacity-building needs regarding the access, use, generation and analysis of digital sequence information on genetic resources

Ethiopia strongly believes that there is a need for technical, financial and human capacity building regarding the access, use, generation and analysis of digital sequence information on genetic resources.

<h3>European Union and its Member States</h3>
--

SUBMISSION

The EU and its Member States are pleased to share their overall views and information on issues related to Notification 2019-012, in line with Decision 14/20 on digital sequence information, adopted at the fourteenth meeting of the Conference of the Parties to the Convention on Biological Diversity.

Pursuant to decision 14/20, paragraph 9, the notification invites “Parties, other Governments, indigenous peoples and local communities, relevant organizations and stakeholders to submit views and information: (a) To clarify the concept, including relevant terminology and scope, of digital sequence information on genetic resources and if and how domestic measures on access and benefit-sharing consider digital sequence information on genetic resources; (b) On benefit-sharing arrangements from commercial and non-commercial use of digital sequence information on genetic resources.”

Pursuant to decision 14/20, paragraph 10, the notification invites “Parties, other Governments and indigenous peoples and local communities to submit information on their capacity-building needs regarding the access, use, generation and analysis of digital sequence information on genetic resources, in particular for the three objectives of the Convention.”

While recalling that digital sequence information (DSI) is not defined and the term is used as a placeholder, the EU and its Member States recognise the complexity of the issue and the need to deepen the understanding of what situations and activities the term DSI might refer to. In our view, DSI is a broad issue with potential implications for the three objectives of the CBD. Hence, clarification of the concept and scope of DSI and the activities it is associated with should be carried out in a comprehensive manner, taking into account the broader CBD context and its three objectives.

The EU and its Member States emphasize the importance of adopting a science-based approach when exploring a possible common understanding of the scope, definition and concept of DSI. The process set up under decision 14/20 para 11 – namely the four studies and the establishment of an ATHEG - should provide such a science-based approach and assist in further clarifying the concept and the scope of DSI, as well as assist with the identification of possible appropriate terminology. We believe that it is important not to pre-empt the results of the scientific studies commissioned, as well as the outcomes of the AHTEG. That being said, the EU and its Member States would like to underline the following aspects with respect to decision 14/20 para 9, point (a) and (b).

A definition of DSI should be based on appropriate relevant terminology and clear terms. In our view, general terms, such as “information”, that might be open to interpretation, do not bring further clarity to the concept and should be avoided. To identify appropriate and more precise terminology, it can be useful to explore the terms commonly used by the scientific community in the context of genetic research. These terms include, for instance, genetic sequence data, nucleotide sequence data, and genetic sequences. As pointed out in the Fact-finding study on DSI (CBD/DSI/AHTEG/2018/1/3), differences in scientific terminology reflect differences in the material referred to, as well as the speed and transformative nature of technological change of today. We are aware that this may constitute an additional challenge to the identification of harmonized terminology.

We note that relevant terminology is used also in international organizations in the context of their activities, such as the WHO. In particular, the WHO PIP Framework uses the term genetic sequence data, and defines genetic sequences as: “The order of nucleotides found in a molecule of DNA or RNA. They contain the genetic information that determines the biological characteristics of an organism or a virus”. It would be useful to take into account relevant existing and agreed terms and definitions for the eventual establishment of a clear terminology for DSI.

The EU and its Member States consider that DSI is not equivalent to a genetic resource. We consider that within the framework of the CBD and the Nagoya Protocol the access to DSI, held in digital or in any other form, is not equivalent to access to genetic resources from which it is generated. In other words, PIC cannot and should not be required for access to DSI, including from publicly available databases.

That said, we note that conditions for generating and using DSI (in case of non-commercial as well as commercial use), which come from utilization of genetic resources within the scope of the Protocol, can be specified in mutually agreed terms (MAT) when a genetic resource is accessed in accordance with domestic measures on access and benefit-sharing. For instance, a permit granted for non-commercial use of genetic resources may envisage that the result of such utilization, including from activities producing DSI, are subject to the conditions set in that permit and can be limited to non-commercial purposes or set further conditions in case of future commercial purposes.

Generation, access to and use of DSI may have important and positive effects on the conservation and sustainable use of biological diversity. The sequenced data are the basis of a large number of biodiversity-related research activities, such as taxonomic research. Furthermore, DSI is widely used for *in-situ* and *ex-situ* conservation purposes, monitoring and implementation of protected areas and monitoring of invasive alien species.

We also stress that DSI is very important for protection of human, animal and plant health. The timely and unrestricted access to such data is crucial for the fast and effective reaction to threats to public health and for routine tasks in health protection. Examples of activities where DSI plays a crucial role include tracing of epidemic outbreaks, vaccine development, tackling antimicrobial resistance, food safety and surveillance and control of infectious diseases; these outcomes can be considered to be a form of non-monetary benefit sharing.

We also underline the importance of DSI in research and innovation and recognize that DSI is also used by other sectors. While fully respecting the rights of countries, we believe that accessibility to data for research purposes should not be subject to heavy administrative measures that could slow down the information sharing and may hinder further increase in knowledge.

We emphasize that public or open access databases are maintained by public funding, and data published in these databases are the results of researchers' work made available to freely share information, including DSI. Bearing this in mind, public databases and open access to their data are a form of non-monetary benefit sharing and contribute to the fair and equitable sharing of benefits.

With regard to decision 14/20, para 10, the EU and its Member States acknowledge that many countries may lack capacity to generate, access and use digital sequence information on genetic resources. Lack of capacity in this regard can hinder the successful implementation of the three objectives of the CBD.

The EU and its Member States strongly support the development, maintenance and increase of capacity in all States to generate, access and use DSI. In this context, we underline that open access to data and free circulation of information available in public databases, as well as the development and publication of tools to analyse the data, are at a basis of several education and scientific research projects, training and other activities carried out in the framework of development and cooperation by the EU and its Member States with other Parties. In effect, the EU and its MS are actively involved in research cooperation projects involving the study of genetic resources with many countries and their scientific institutions. This cooperation also contributes to build capacity building in relation to DSI in those countries and their research entities. That said, we are aware that additional capacities and appropriate technology may be needed to use DSI and we stand ready to further engage with the other Parties to consider the capacity-building needs and the appropriate means to support this.

India

SUBMISSION

- a) To clarify the concept, including relevant terminology and scope, of digital sequence information on genetic resources and if and how domestic measures on access and benefit-sharing consider digital sequence information on genetic resources;
- (b) On benefit-sharing arrangements from commercial and non-commercial use of digital sequence information on genetic resources.

Concepts Scope and terminology

Genetic sequence information is about the manner in which the base pairs of DNA and RNA are structured in an organism. Digital sequence information (DSI) on genetic resources (GRs) may include any information derived out of these nucleotide or amino acid sequences, their biochemical composition, gene locations, genetic maps, chromosome maps, localization details, functional expression details, genome sequences and their annotations, and applied aspects in decoding the sequence information. The scope of DSI thus would inter alia cover information on genetic or biochemical composition including nucleic acid sequence gene reads and associated data, sequence assembly, its annotation and gene mapping describing the genome, individual genes or fragments thereof, organelle genome, gene expressions, macromolecules and cellular metabolites.

The basic difference between the information content in tangible genetic resources and the digital sequence data lies in the fact that the digital format, as opposed to the genetic expression contained in the tangible

genetic resource, converts the genetic information into a public good without changing the content as such. The digitised format makes it easily accessible and transferable.

Genetic sequence information in digitised format is often publicly available and accessible through various genome databases. With advanced genetic sequencing and synthesising technologies, the digitised genetic sequence information may preclude the need to access any genetic material in tangible format or physical access of the biological material.

As regards terminology, DSI is not a very appropriate term for the different kinds of information on genetic resources. The terminology therefore requires further consideration. Sequence Information, or Sequence Data are some suggestions, that merit consideration.

Benefit sharing arrangements from use of DSI on GRs

Rapid technological advancements have resulted in emergence and growth of digital sequence information as a faster, cheaper and more accurate tool, the use of which offers opportunities as well as challenges for realizing benefits.

There is a general recognition of positive potential role that DSI on GRs may have in conservation and sustainable use of biodiversity (i.e., the first two objectives of CBD). The fact that DSI enables access to genetic information of an organism without physically accessing it, directs much of the focus of discussions towards impact of DSI on sharing of benefits arising from utilization of GRs, which constitutes the third objective of CBD, operationalised through the Nagoya Protocol, while acknowledging that the three objectives of the Convention are interlinked and mutually supportive.

CBD defines genetic resource as “genetic material of actual or potential value” and genetic material is defined as “any material of plant, animal, microbial or other origin containing functional units of heredity”. It is clear from these two definitions that the CBD as well as the Nagoya Protocol recognise both the tangible and intangible elements in the genetic material, i.e., the physical material as well as the value it contains. The value of the genetic material lies in the genetic information it contains, whether in the actual form or potential form.

The definition of utilization of genetic resources under the Nagoya Protocol reads as “research and development on the genetic and/or the biochemical composition of genetic resources, including through the application of biotechnology”. As per this definition, utilization of genetic resources is not confined to R&D on the tangible genetic resources, but is extended to activities over the genetic and biochemical composition of such resources. Genetic as well as biochemical composition is nothing but the gene sequence whether digital or tangible. Synthesising the DNA with access from digital sequence would therefore fall within the scope of utilization under the Nagoya Protocol. In other words, accessing digital sequence information amounts to accessing the genetic resource itself, and its utilization would fall within the scope of CBD/Nagoya Protocol, qualifying for application of ABS regulatory framework even though there is no physical access of the genetic material. While for non-commercial research, simplified procedures may be considered, commercial utilisation would trigger benefit sharing as per the provisions of CBD/ Nagoya Protocol.

This argument is also supported by the non-exhaustive list of different uses of genetic resources by the CBD’s Group of Legal and Technical Experts on Concepts, Terms, Working Definitions and Sectoral Approaches which formed the basis of the definition of utilization of genetic resources in the Protocol. This list expressly referred to the synthesis of DNA segments.

As was also concluded in the well-researched fact finding study commissioned by the Secretariat, the challenge in the context of ABS related to identification of contributors, users and provenance; and monitoring the utilization, including how to trace commercialization of use of DSI especially which is freely

available in public database can be overcome with the use of technological tools, in line with the relevant provisions of the CBD and Nagoya Protocol. This is amply demonstrated by some of the examples that are given in that study (such as GSAID, use of unique identifiers, GGBN, Oldham's model etc).

How domestic measures on ABS consider DSI on GRs

In India, the ABS provisions of CBD and the Nagoya Protocol on ABS are implemented inter alia through the Biological Diversity Act, and regulations thereunder. While these do not include explicit reference to DSI or any such terminology, the relevant provisions in the Act can cover in their scope the utilization of DSI. For example, the term research as defined in Section 2 would cover DSI. Similarly, the requirement prescribed in Section 6 which refers to 'information on biological resource' would cover DSI.

Capacity building needs

There is a need for enhancing capacity to generate, analyse, access and use digital sequence information on genetic resources.

Iran

SUBMISSION

Reference to Notification 2019-12, dated 5 February 2019, I wish to submit Islamic Republic of Iran's view on "Digital Sequence Information on Genetic Resources (DSIGR)":

Rapid development of modern Biotechnology and molecular techniques including identification, isolation, modification and successful transfer of genes and/or any DNA controlling gene function in addition to recent development of techniques for genome editing and advances in data analysis has increased the potential value of "Genetic Resources" in one hand and attempts to access these resources and/or to their digital sequence information. Therefore, a science and policy- based process on digital sequence information on genetic resources has been adopted in the 14th MOP.

The scope of Digital Sequence Information on Genetic Resources should be extended to any "sequence information originated from genetic resources" and should be included all DNA/RNA and amino acid sequences. This information should also include both raw sequences information or the processed/annotated sequence information. This information should be made available and accessed by public for non-commercial use only. However, any commercial application of the Digital Sequence Information on Genetic Resources should be followed the same procedures and terms decided on Nagoya Protocol on access and benefit sharing.

In other words, there should be no difference in accessing and commercial application of genetic resources as stated in Nagoya Protocol and the Digital Sequence Information of the same genetic resources.

By non-commercial use, we mean access and use of Digital Sequence Information on Genetic Resources for educational/training and/or for research purpose only.

In spite of its significance, developing countries lack infrastructure for generation of Digital Sequence Information from their own genetic resources, storing, the generated sequence and their analysis. Owner of technologies should make technologies available for developing countries with no restriction (including politically motivated economic sanctions).

Capacity building requires provision of access to hardware (equipment/machinery) and software (computer programs and know how), for the three objectives of the convention.

Iran wishes to be actively involved in extended AHTEG and to contribute for a science and policy- based process on digital sequence.

Japan

SUBMISSION

In response to Notification 2019-012, Japan would like to submit its views as follows:

(1) Views and information to clarify the concept, including relevant terminology and scope, of digital sequence information on genetic resources and if and how domestic measures on access and benefit-sharing consider digital sequence information on genetic resources (decision 14/20, paragraph 9 (a))

(i) Terminology

Japan believes that the term “genetic sequence data” (GSD) is most appropriate for discussion under the Convention for the reasons stated below.

- The term GSD is widely used by the scientific community, thus entailing scientific validity. Furthermore, GSD is an established term under the World Health Organization’s (WHO) Pandemic Influenza Preparedness (PIP) Framework.
- GSD imposes clarity on the scope of the terminology, thus easily lending itself to a common understanding among Parties, helping to remove ambiguity within the discussions. We can refer to the definition and use of the term “genetic sequences” under the WHO PIP Framework --- which states that “the order of nucleotides found in a molecule of DNA or RNA. They contain the genetic information that determines the biological characteristics of an organism or a virus”.
- The term GSD is consistent within the context of paragraph 4 and paragraph 7 of decision 14/20. Those paragraphs indicate that, in most cases, digital sequence information (DSI) on genetic resources is generated during the process of, or, as a result of the utilization of genetic resources accessed. GSD is, in most cases, generated through determination of nucleotide sequence of genetic resources that is regarded as a case of “research and development on the genetic and/or biochemical composition of genetic resources” (Article 2 (c) of the Nagoya Protocol).

(ii) Scope

The discussion on DSI/GSD should not affect the physical and temporal scope of the Convention. In addition, it is a matter of course that DSI/GSD on human genetic resources is excluded from the scope of the discussion, since it was reaffirmed at COP2 that human genetic resources are not included within the framework of the Convention.

Likewise, the scope of the terminology should also exclude matters that have already been or supposed to be addressed by other intergovernmental organizations, such as the WHO and the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA).

(iii) If and how domestic measures on access and benefit-sharing consider digital sequence information on genetic resources

Japan does not take measures on access and benefit-sharing as a provider country of genetic resources. However, Japan harbors the following views on this topic, based on its legal interpretation of the Convention and the Nagoya Protocol:

- “Genetic resources” defined in Article 2 of the Convention refer to tangible materials, and thus do not include DSI/GSD and any other information.
- As such, access to DSI/GSD is not subject to prior informed consent (PIC).

(2) Views and information on benefit-sharing arrangements from commercial and non-commercial use of digital sequence information on genetic resources (decision 14/20, paragraph 9 (b))

Mutually agreed terms (MATs), concluded at the time of access to a genetic resource for its utilization, may theoretically cover benefit-sharing from the use of DSI/GSD on the genetic resource, as written in paragraph 7 of decision 14/20. However, given that MATs are negotiated between a provider and user of the genetic resource, their content is usually confidential to other parties. As such, Japan does not have any specific insights/ information on benefit-sharing arrangements resulting from the use of DSI/GSD.

Nonetheless, Japan believes that open access to DSI/GSD is a form of benefit-sharing, as it actually contributes to the conservation and sustainable use of biodiversity as well as to other important areas such as food security and human/ animal/ plant health (see the previous submission from Japan in September, 2017: <https://www.cbd.int/abs/DSI-views/JAPANDSI.pdf>). We should avoid taking the discussion on DSI/GSD toward a direction that may hinder such benefits from an open access to DSI/GSD.

We should also discuss what “the use of DSI/GSD” actually refers to before considering this matter further.

(3) Information on capacity-building needs regarding the access, use, generation and analysis of digital sequence information on genetic resources, in particular for the three objectives of the Convention (decision 14/20, paragraph 10)

From Japan’s experiences of supporting developing countries, we perceive the needs for capacity-building in the following areas, among others:

- the use of DSI/GSD for the analysis of environmental DNA (DNA samples of organisms and their parts extracted from environmental samples such as soil and water) for inter alia identifying and monitoring the species composition in an environment;
- the generation and use of DSI/GSD for taxonomical works, whose capacity is needed when establishing a culture collection of microorganisms, for example.

We also believe ensuring open access to DSI/GSD is essential in capacity-building efforts.

Madagascar

SUBMISSION

We are pleased to submit our view on Digital Sequence Information (DSI).

DSI should be considered as derived from genetic resources, so they fall within the scope of the Nagoya Protocol. Indeed, without genetic resources, there will be no DSI.

It is therefore necessary to integrate into the ABS national regulation, that as soon as the research relates to genetic studies, clear and precise devices are essential on the use of the DSI, their capitalization in databases or their transfer to third parties.

In particular, the MAT that will be established for access to resources, for which a genetic study is being considered, must specify that the DSI thus obtained derived from which resource from which country(ies).

It is also conceivable to design a specific permit for the access and use of DSI.

It is, of course, necessary to have capacity building in the field of DSI to establish these national regulations. The main principles should be shared between supplier countries, especially megadivers ones (in Africa, Asia, Latin America, etc.) because the resources used to generate DSI could come from different origins.

Mexico

SUBMISSION

- a) Aclarar el concepto de información digital de secuencias de recursos genéticos, incluyendo terminología relevante y alcance; y de ser el caso, aclarar de qué maneras las medidas domésticas sobre acceso y reparto de beneficios consideran la información digital de secuencias de recursos genéticos.

Contexto general de la “información digital sobre secuencias” y la biodiversidad.

El acceso a la información de secuencias genéticas de acceso libre (gratuito) es cada vez mayor y ha reducido la necesidad de requerir muestras de material biológico (plantas, tejido animal, microorganismos) de manera física debido a diversas situaciones tales como: el progreso tecnológico en materia de Biología molecular; de secuenciación de grandes fragmentos de ADN a un menor precio; la posibilidad de editar genomas; a la creciente capacidad informática para el almacenamiento y la transferencia de datos; a la posibilidad de sintetizar secuencias de genomas mayores a un menor costo; el interés en la generación de repositorios o bases de datos, e incluso la tendencia económica a valorizar y mercantilizar la información.

Asimismo, la disponibilidad de secuencias genéticas y su acceso irrestricto, podría volver obsoleto el tercer objetivo del Convenio sobre Diversidad Biológica: el reparto justo y equitativo de los beneficios derivados de la utilización de los recursos genéticos; así como los esfuerzos de implementación del Protocolo de Nagoya sobre acceso a los recursos genéticos y la participación justa y equitativa de los beneficios, dado que los marcos jurídicos de los países no han sido actualizados para incluir las nuevas realidades técnicas, y económicas.

Concepto de información digital de secuencias de recursos genéticos.

En la Decisión 14/20, adoptada por la Conferencia de las Partes del Convenio sobre la Diversidad Biológica en la decimocuarta reunión celebrada en Sharm el-Sheikh, referente a la información digital sobre secuencias de recursos genéticos, la Conferencia de las Partes observó que el término “información digital sobre secuencias” podría no ser el más apropiado, y que este se utiliza de modo provisional en tanto se acuerde un término alternativo. Asimismo, se reconoció la importancia de la información digital sobre secuencias de recursos genéticos para los tres objetivos del Convenio de Diversidad Biológica que se apoyan mutuamente.

A este respecto, la Secretaría Ejecutiva de la CIBIOGEM considera que, si bien el concepto “información digital sobre secuencias” es un término empleado hasta ahora a manera de “vehículo” para las sesiones de trabajo y discusión al seno de las reuniones de la Conferencia de las Partes del Convenio de Diversidad Biológica, los procesos y consecuencias que se pudieran generar sobre el tema son aún más importantes que las palabras en sí; por lo que la definición del concepto es solamente una parte de la aproximación a este asunto.

Consideramos que la discusión al seno del Convenio de Diversidad Biológica, del Protocolo de Cartagena, y del Protocolo de Nagoya, debe centrarse en los tipos de información sobre los recursos genéticos vegetales para la alimentación y la agricultura, y en general, los cuales deben necesariamente incluir a las secuencias de Ácido desoxirribonucleico (ADN), incluido el cADN; el Ácido ribonucleico (ARN) en todas sus formas;

secuencias de amino ácidos; “SNPs”; “STRs”; todo el ámbito de información molecular (por ejemplo, patrones de metilación del ADN); así como meta datos asociados.

Tal y como lo han señalado algunos participantes del Grupo de Expertos Técnicos Ad hoc (AHTEG) en Información digital sobre secuencias del Protocolo de Cartagena, coincidimos en que el uso de la palabra “digital”, podría ser confuso. Aunque se comprende que la palabra “digital” se usó en el sentido de referencia a la velocidad y facilidad en que los nuevos desarrollos tecnológicos permiten la transferencia de secuencias entre fronteras, de manera frecuente, pero no exclusivamente por internet (medio digital). Probablemente no debería incluirse en la definición final del concepto, ya que continuamente existen adelantos tecnológicos que permiten el almacenaje y transferencia de información (por ejemplo, la computación cuántica), o incluso, implica el uso -aún vigente- de métodos, procedimientos o herramientas de almacenaje de información que no son digitales, tales como el papel simple.

Por otro lado, actualmente se han identificado diversos actores clave en el tema de la “información digital sobre secuencias”, que se posicionan de acuerdo a sus necesidades y potenciales beneficios. Algunos de estos actores son: las bases de datos privadas, las bases de datos públicas, los pueblos indígenas y comunidades locales. Se explican a continuación características generales de dichos actores clave.

Bases de datos privadas: Se puede suponer que el objetivo principal de las bases de datos privadas es el desarrollo de productos, así como el generar propiedad intelectual de valor agregado. Algunas compañías semilleras han generado grandes bases de datos de recursos genéticos vegetales para la alimentación y la agricultura, lo cual, como se mencionó antes, podría poner en riesgo los procesos y mecanismos contemplados para el reparto justo y equitativo de los beneficios derivados de la utilización de los recursos genéticos y a los mismos recursos genéticos.

Bases de datos públicas: Las bases de datos públicas son de gran relevancia debido a que pueden tener un enfoque académico o de uso y beneficio social. El intercambio y uso de la información digital de las secuencias es una tarea esencial de la investigación científica. Existen bases de datos públicas de gran tamaño, tales como la base de datos conocida como Genbank o la Base de Datos llamada Archivo europeo de nucleótidos (ENA, por sus siglas en inglés). Debe considerarse que el hecho de que sean de acceso público no debería ser tomado como sinónimo de que sean “gratis”, o de “libre acceso”. Actualmente muchas bases de datos públicas no han hecho manifestaciones relevantes sobre la manera en la que se obtuvo la información que está puesta a disposición en la base de datos; o si existen restricciones de propiedad intelectual o uso comercial para el uso de la misma, por lo que deben elaborarse mecanismos que aseguren la operatividad de los procesos que garanticen el reparto justo y equitativo de los beneficios, de lo contrario, los gobiernos de los países en desarrollo, los agricultores y los pueblos indígenas que crearon y preservan la biodiversidad serán afectados en el corto plazo.

Pueblos indígenas y comunidades locales: Los pueblos y comunidades indígenas, bajo una gran diversidad de circunstancias han generado y aún mantienen y recrean conocimientos tradicionales asociados a los recursos genéticos que se encuentran en sus territorios, por lo que deben ser reconocidos como los titulares legítimos de esos conocimientos tradicionales asociados a los recursos genéticos que se preservan ya sea de manera oral, documental, o de otra forma. Adicionalmente, el reciente reporte sobre el estado de la biodiversidad y los servicios ecosistémicos, elaborado por parte de la Plataforma para la Biodiversidad y los Servicios ecosistémicos (IPBES), ha documentado la labor que los pueblos indígenas y comunidades locales realizan en la conservación de los recursos genéticos a partir de los conocimientos tradicionales que poseen, por lo que deben afianzarse los mecanismos para asegurar el reparto justo de beneficios, derivado del acceso y uso de los recursos de la diversidad biológica. Tal y como lo menciona la Decisión CBD/COP/DEC/14/17 adoptada por la Conferencia de las Partes del Convenio de sobre la Diversidad Biológica: “ alienta a los Estados Parte a colaborar con las comunidades indígenas y las comunidades locales en la aplicación del Convenio, incluidos los esfuerzos de protección y conservación de sus territorios y zonas...”. A este respecto, podemos mencionar que falta trabajo en la construcción de capacidades técnicas e implementación de infraestructura que permita incorporar a los pueblos indígenas de manera más activa a la discusión y decisiones sobre la información digital sobre secuencias, el acceso y uso de los

recursos de la diversidad biológica, y a los procesos y decisiones para apuntalar los mecanismos que aseguren el reparto justo y equitativo de los beneficios derivados de la utilización de los recursos genéticos.

- a) Acuerdos sobre el reparto de beneficios para el uso comercial y no comercial de la información digital de secuencias de recursos genéticos.

La Secretaría Ejecutiva de la CIBIOGEM actualmente no cuenta con información relativa a acuerdos sobre el reparto de beneficios para el uso comercial y no comercial de la información digital de secuencias de recursos genéticos.

Republic of Korea

SUBMISSION

In response to CBD Notification 2019-012, the Republic of Korea would like to provide its current views and information on Digital Sequence Information (DSI) on genetic resources. At the outset, the Republic of Korea reiterates that the term “DSI” is mere a placeholder as noted in the decision CBD/COP/DEC/14/20.

1. Clarification of the concept of digital sequence information on genetic resources including relevant terminology and scope

The Republic of Korea suggests that “digital” should be removed from the term “digital sequence information”. Biologists do not use the term “digital sequence” when referring to “sequence”. Considering the general usage of the word “digital”: digital technology, digital display, digital camera, etc., “digital” indicates particular technologies or devices dealing with discrete data. All sequences generated from genetic resources can be described, stored, and transferred digitally. As sequences themselves have a digital nature, the word “digital” in the term “digital sequence information” is redundant. Using the more commonly-understood term “genetic sequences” is appropriate as it delimits the scope of “sequences” related to genetic resources and provides clarity in the terminology.

As “information” is neither specific nor limited, “data” is a more appropriate term than “information”. The Oxford English Dictionary defines “data” as “facts and statistics collected together for reference or analysis” and “information” as “facts provided or learned about something or someone”. Therefore the Republic of Korea prefers to use “data” in this discussion.

Inferred by the current discussion on the placeholder, DSI in CBD, the Republic of Korea prefers “genetic sequence data (GSD)”, since GSD is commonly used in the scientific field. However, we would like to reiterate that GSD is not a genetic material or genetic resource. Hereinafter, in this submission, the view and information on domestic measures on access and benefit-sharing and benefit-sharing arrangements from commercial and non-commercial use of DSI was written in considering, instead of DSI, the concept and scope of GSD.

2. If and how domestic measures on access and benefit-sharing consider digital sequence information on genetic resources

There are no domestic restrictive measures on access and benefit-sharing considering genetic sequence data on genetic resources in the Republic of Korea. The Republic of Korea believes that free access to genetic sequence data would promote scientific advancement and contribute to the conservation and sustainable use of biodiversity.

3. On benefit-sharing arrangements from commercial and non-commercial use of digital sequence information on genetic resources

As genetic sequence data is not genetic resources under the Convention on Biological Diversity, it is not subject to the Article 15 on CBD. The Republic of Korea is of a position that the priority should be imposed on building consensus among the Parties on whether GSD is subject to access and benefit-sharing prior to further discussion.

South Africa

SUBMISSION

- (a) To clarify the concept, including relevant terminology and scope, of digital sequence information on genetic resources (DSI) and if and how domestic measures on access and benefit-sharing consider digital sequence information on genetic resources**

DSI has a significant role to play in environmental, life sciences and biodiversity research such as population, pathogenicity and epidemiological studies. The study results thereof are used to protect human, animal and plant health, food security and ecosystem functioning. DSI is already widely shared for the surveillance of infectious diseases (in humans, animals and crops) and invasive species across the globe. There is a lot of DNA sequence data generated mostly from research on species / strains and populations of plants, animals, fungi, bacteria and viruses. Currently the generation of such sequences are part of academic research that results in publication in peer-reviewed journals. Most journals require that the sequences included in a paper are published in a globally accessible repository such as NCBI / Genbank. In South Africa, funding programmes such as the Foundational Biodiversity Information Programme (FBIP), which is funded through the Department of Science & Technology, supports the generation of DNA barcodes which are sequences of specific genes that will allow the identification of material to species level for plants, animals and fungi. Barcode sequences are submitted to the Barcode of Life Database (BOLD) which is an open global repository that includes the sequence data, but also data about the actual specimens from which the tissue / sample was taken to generate the sequence data. Researchers, and enforcement agencies use the sequence data from the global community that is made accessible in BOLD and / or NCBI for comparison purposes and for identifying materials. South Africa, through the Department of Science & Technology, has signed an MoU with the International Barcode of Life (IBOL) and this promotes contributing sequences to the global repository and making those openly accessible through the online repository, BOLD. However, there is no direct benefit-sharing from the use of /access to the sequence data with the country of origin (i.e South Africa).

As noted by the CBD-COP decision 14/20 jointly with COP-MOP decision NP-3/12 that there is an understanding that the term “**digital sequence information**” may not be the most appropriate term and that it should be used as a placeholder until an alternative term is agreed. Therefore, the debate on terminology should not be used to delay progress on this matter. We need to consider what constitute the genetic resources and work on a definition that explicitly make reference to the genes or genetic material which constitute a genetic resource. There are, of course, other possible terms that could be used such as genetic information, natural information, genetic sequence data, and digital sequence data.

The 2013 amendment to the National Environmental Management: Biodiversity Act, 2004 (Act No. 10 of 2004) (NEMBA) ensures that the nation’s indigenous genetic and biological resources are developed and utilized in an ecologically sustainable manner while promoting social and economic development in particular in the areas where the indigenous genetic or biological resources and associated traditional knowledge is accessed, as one of its intentions. These amendments included the following definitions which are linked to DSI: “**Derivative**”, in relation to an animal, plant or other organism, means any part, tissue or extract of an animal, plant or other organism, whether fresh, preserved or processed, and includes any genetic material or chemical compound derived from such part, tissue or extract; “**Genetic material**”, means any material of animal, plant, microbial or other biological origin containing functional units of heredity; “**Genetic resource**”, includes any genetic material; or the genetic potential, characteristics or information of any species. These resources may be gathered from the wild or accessed from any other source. Prior Informed Consent and establishment of the Mutually Agreed Terms are compulsory

requirements for requesting commercial access to indigenous genetic and biological resources and also for access to traditional knowledge associated with indigenous genetic and biological resources. South Africa believes that creation of sequence information from genetic resources depends on the physical access to the genetic resources. Therefore, the Mutually Agreed Terms and the permit templates included in the NEMBA, provides mandatory clauses that addresses third party transfer terms and conditions which could include the utilization of the DSI on genetic resources, whether stored in the public or private databases.

It is clear from the South Africa's example that the Parties to the Nagoya Protocol on ABS should continue their work towards finding innovative/creative international policy solutions aimed at ensuring fair and equitable sharing of benefits with the country of origin where the original genetic resources contributed in the creation of sequence data which have been accessed and commercialized.

(b) On benefit-sharing arrangements from commercial and non-commercial use of digital sequence information on genetic resources

In terms of non-commercial use of sequence data, the repositories are considered to be global resources that are both contributed to and used by a very wide and large community. This is critical to ensure that research is incremental, cost effective and efficient (rather than each study having to redo all sequencing, only new materials needs to be sequenced and existing information can be accessed and used). The benefit-sharing is probably not made explicit on the websites of the repository and could be considered to be non-monetary. There is an unwritten sentiment that if a large amount of sequence data from a single researcher is being used by another researcher, then there should be a collaboration and co-authorship, but this does not always happen. There are also instances of someone trawling repositories such as BOLD to identify new species in the database, and then describing these and benefitting from the publication without involving the researcher who collected the species and generated the sequence data. It is known that currently international researchers collect / sample South Africa species, and take the materials out of the country, and there is no national benefit from such international research work and it is very difficult to trace the benefits that are being generated without an internationally agreed standard.

In terms of the extent of commercial use of the DSI, South Africa does not have concrete case study, largely due to the fact that DSI might be accessed under academic research terms, uploaded onto databases, and end up being used commercially, potentially by multiple different users, without the original providers aware of or involved in this process.

In terms of the Human Biological Materials, the South African government through the Department of Health has gazetted a Material Transfer Agreement in 2018 which applies to all the providers and recipients of the human biological materials for use in commercial or non-commercial research. Importantly, the definition of “**materials**” provided include data. This agreement seek to ensure that the benefits derived from commercial or non-commercial research are shared in a manner that is fair and equitable. The scope of benefits includes, amongst others, the sharing of information, use of research results, royalties, acknowledgement of the Provider as the source of the materials, publication rights, transfer of technology or materials, and capacity building. This agreement recognises secondary use of materials which could be in a form of DSI. This agreement also emphasise the fact that ownership of the human biological materials remains with the country of origin regardless of where such data and their format is stored, until they get destroyed.

(c) On their capacity-building needs regarding the access, use, generation and analysis of digital sequence information on genetic resources, in particular for the three objectives of the Convention

South Africa is mindful that the CBD Secretariat has already conducted a series of capacity building trainings as part of its Global Taxonomy Initiative (GTI). The main focus for these training was largely on DNA Barcoding and the targeted participants were cross-border regulatory authorities, forestry/fisheries authorities and protected area managers. These efforts were mainly to promote conservation of biodiversity and sustainable use of its components. However, the socio-economic development aspect thereof still

requires attention in the context of fair and equitable sharing of benefits derived from the commercial utilization of DSI.

Considering that South Africa is a Party to the Nagoya Protocol on ABS and also a key negotiator on DSI, and also that it has a number of DSI related initiatives at national level, the following have been identified as the Country's areas for capacity development:

- Scientific / Academic Institutions, especially researchers and technical staff who collect and process and manage biological/ genetic materials that are used for DNA sequencing requires capacity building on global standards including the Nagoya Protocol on ABS. Although, South Africa is running various training sessions through the FBIP on using BOLD which target researchers and students, but more training is needed. There has been training of enforcement staff and prosecutors on understanding what DNA barcodes are and how they can be used because the technology is relevant to illegal trade, but more training for the biosecurity and bioprospecting sectors is needed in the context of sequence data.
- Regulatory institutions on environment, science and technology, trade and industry, agriculture, and health who are responsible for developing and implementing strategies, policies and legislations relating to Access to Genetic Resources and the Fair and Equitable Sharing of Benefit Arising from their Utilisation in order to understand the rapid technological advancement in the space of access and utilization of genetic resources towards DSI, so that they can adapt their legal instruments accordingly.
- To generate own DSI for commercial exploitation purpose in a form of sequencing equipments/training/capabilities and foundry support for synthesizing modified/synthetic sequences for R&D and economic development purposes. Although, the Department of Science & Technology is funding two large national research infrastructure initiatives – the Natural Science Collections Facility and the National Biodiversity Biobanks, which aims to ensuring that there are materials representing South Africa's biodiversity for use by researchers and also for commercial use, additional support in this regard is therefore needed.

NOTE: Compiled by the National Focal Point: Nagoya Protocol on ABS

Switzerland

SUBMISSION

Switzerland thanks the Executive Secretary for the Notification 2019-012 inviting for views and relevant information regarding the specific aspects on digital sequence information on genetic resources in accordance with decisions CBD 14/20 and NP-3/12. This submission represents the current reflections of Switzerland regarding the points raised in notification 2019-012, but does not prejudice the Swiss position in the upcoming negotiations.

1. General comments

The science- and policy-based process on digital sequence information on genetic resources (in the following “DSI”) launched by decisions CBD 14/20 and NP-3/12 will be an important contribution to the common understanding of the concept of DSI and its potential implications on the objectives of the Convention on Biological Diversity and 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. As stated in the submission on DSI by Switzerland in 2017 (<https://www.cbd.int/abs/DSIviews/Switzerland-DSI.pdf>), we are of the opinion that clarifying the terminology related to DSI is an important first step for the ongoing discussions on this topic in different fora.

Technological developments, such as the increased generation, availability and use of DSI, may be highly relevant for the conservation and sustainable use of biological diversity, and hence to achieve the first two

objectives of the Convention. In the context of the third objective of the Convention and the objective of the Nagoya Protocol, a deeper understanding of the relevance of DSI for access to genetic resources and the fair and equitable sharing of benefits arising from their utilization is required. While the discussions on DSI are ongoing, particular caution is to be exercised not to take steps that run counter to these objectives, such as decisions, which may hinder the characterization and documentation of biological diversity, including of genetic resources. As a first priority, the implementation of the Convention should be promoted and strengthened, as well as the implementation of the Nagoya Protocol and its ratification by all Parties of the Convention and other countries.

2. Responses to the specific points raised pursuant to decision 14/20, paragraph 9

(a) To clarify the concept, including relevant terminology and scope, of digital sequence information on genetic resources and if and how domestic measures on access and benefit-sharing consider digital sequence information on genetic resources;

Clarification of the concept of DSI, including relevant terminology and scope:

The discussions on DSI under the Convention and the Nagoya Protocol as well as in other international fora are challenging due to the general lack of consensus or definition of what technically constitutes DSI. Therefore, we are of the opinion that a clarification of the concept and terminology as well as an operational definition of DSI form an important basis for further discussions. Taking note of other fora's discussions on the scope and definition of DSI is a further important element for clarifying the concept.

The report of the AHTEG (CBD/DSI/AHTEG/2018/1/4) lists several different types of information that the term DSI may refer to or that may be relevant in this context. The broad range of different types of information reflects the complexity of this issue. If Parties aim to define a concept for DSI in order to obtain more clarity on potential implications in the framework of the Convention, the concept of DSI must refer to a clearly defined range of information and processes to obtaining, proceeding and transferring that information. Furthermore, it is important to note that not all information listed in the report of the AHTEG necessarily represents information, which is resulting from the utilization of a genetic resource as defined in the Nagoya Protocol (see AHTEG report paragraphs 2 and 3, information (e) to (i)).

For further clarification of the terminology of DSI, the use of the alternative term "genetic sequence data" (GSD), as suggested by some Parties in the Commission on Genetic Resources for Food and Agriculture (see CGRFA-17/19/4/Inf.1) and used within the WHO PIP Framework (Chapter 5.2), should be considered, as well as other potential alternative terms. Those terms also need to be assessed regarding clarity of their concept and scope.

Furthermore, DSI may have several properties and characteristics that are distinct from those of genetic resources (e.g. they are intangible, they may be modified, they are often used in bulk, etc.) and potential provisions on DSI may affect other fields of law besides environmental law (e.g. data protection, intellectual property rights, trade secrets, etc.). Therefore, it should also be carefully examined, which work related to DSI lays primarily within the scope and mandate of the Convention.

While there is a need for clarification of the concept and terminology of DSI, in the view of Switzerland it is essential to be consistent with the use of terms as defined in Article 2 CBD and Article 2 Nagoya Protocol, and other relevant Articles, as they represent the state of the negotiated outcome, which form the base on which Switzerland has ratified these instruments. In the context of the current discussions, the definitions of the terms "genetic resources" as "genetic material of actual or potential value" and "genetic material" defined as "any material of plant, animal, microbial or other origin containing functional units of heredity" are of particular relevance. In the view of Switzerland, the terms "genetic resources" and "genetic material" clearly refer to tangible matter, while DSI does not fulfil the criteria of the definitions of either "genetic material" or "genetic resources" under the Convention and the Nagoya Protocol, and therefore is not covered by those instruments. In addition, the use of these terms may be of relevance for other international instruments dealing with genetic resources.

Switzerland is well aware that the generation, availability and use of DSI is rapidly increasing, and takes note of the current ambiguity of the term and concept of DSI, which indeed offers different interpretations.

Therefore, Switzerland is supportive of the development of a comprehensive and fact-based overview of the technical concept of DSI in different fora as well as of an in-depth understanding of the current respective legal situations in order to contribute to the further discussions on this topic.

Domestic measures on access and benefit-sharing considering DSI:

The access and benefit-sharing legislation in Switzerland relies on the definitions of terms in the Convention and the Nagoya Protocol at the time of ratification. Accordingly, domestic measures in Switzerland do not explicitly address DSI, as these terms and definitions do not cover DSI. Nevertheless, as stated in the previous submission on DSI by Switzerland (<https://www.cbd.int/abs/DSIviews/Switzerland-DSI.pdf>), research on the genetic or biochemical composition of genetic resources can generate different forms of DSI and hence, such information can play a role during the process of utilization of genetic resources. The specific conditions under which a particular genetic resource can be utilized, can be negotiated and defined in the mutually agreed terms (MAT) between the country providing a genetic resource and the user at the time of access to this genetic resource. Swiss contract law indeed provides that the parties may freely determine the terms of a contract. This contract may thus also include partner's mutually agreed provisions regarding DSI resulting from the utilization of the particular genetic resource (e.g. provisions concerning the publication of digital sequence information on the genetic resource). Hence, the use of DSI as well as benefits arising from its use can be covered by MAT (in line with paragraph 7 of decision CBD/COP/DEC/14/20).

(b) On benefit-sharing arrangements from commercial and non-commercial use of digital sequence information on genetic resources.

In Switzerland, users of genetic resources have to submit a notification of compliance with the due diligence requirement to the National Competent Authority before market authorization or commercialization of products developed on the basis of utilized genetic resources. However, as the domestic measures in Switzerland do not explicitly address DSI, we do not have any specific information on benefit-sharing arrangements from commercial and non-commercial use of DSI.

B. SUBMISSION FROM A NON-PARTY

United States of America

SUBMISSION

Terminology and Scope

The United States defines digital sequence information (DSI) on genetic resources to refer to the genetic sequence data (GSD) that describe the order of nucleotides in DNA or RNA in genetic material. Recognizing that there is a lack of understanding of the term DSI on genetic resources, we use the more precise and broadly understood term, GSD, in our response going forward.

We note that GSD are neither genetic material nor a genetic resource. It is essential to maintain a conceptual and definitional distinction between genetic material itself and data associated with that material.

U.S. views regarding access to GSD and benefit-sharing

The United States considers that GSD are essential for scientific research, including research that contributes directly to conservation and sustainable use of biodiversity. It is our view that ready access to publicly available information on genetic resources, including GSD, spurs innovation and provides other broad benefits for society, including enhancing scientific research and collaboration, increasing food security, and protecting public health.

We believe that the scientific norm of rapidly sharing information, including GSD, fosters international collaboration and is itself a form of benefit-sharing that creates other non-monetary benefits, such as

voluntary capacity-building, education, and training. Recognizing that, the United States has made it a policy priority to promote access to scientific data resulting from federally funded research. We consider that the greater the amount of GSD that are shared broadly for study and comparative analysis, the greater the potential benefit for society.

As best practice, GSD are publicly available via international data repositories, such as GenBank and the International Nucleotide Sequence Database Collaboration, as well as in journals found in print and online, including pre-print publication resources. Most major scientific journals require that supporting data, such as GSD, be submitted to a publicly accessible database in order for manuscripts to be published. Both the U.S. National Institutes of Health and the National Science Foundation require papers that contain federally funded research data be made publicly accessible within 12 months of publication.

The U.S. Government supports several initiatives that enable the timely exchange of data that are essential for research, innovation, and protecting human, animal, and plant health. For example, the United States is a party to the FAO International Treaty on Plant Genetic Resources for Food and Agriculture, a specialized international ABS instrument that explicitly recognizes exchange of information as a mechanism for benefit-sharing (ITPGRFA Article 13.2). The United States is also a partner in the Global Open Data for Agriculture and Nutrition (GODAN) initiative, which promotes proactive sharing of open data to address challenges and advance food security. In addition, U.S. researchers are collaborating internationally on the Earth BioGenome Project, an effort to sequence and make available GSD about all eukaryotes.

We believe that regulations that restrict or delay access to and sharing of GSD would likely lead to a significant reduction in data-sharing, which could negatively impact efforts to promote the sustainable use and conservation of biodiversity. We consider that moves to force changes to procedures for information management within laboratories would also carry significant costs and have negative implications for innovation. In our view, these dynamics could stifle research, hindering progress in agriculture, human and animal health, and other sectors.

We maintain that timely access to data, including GSD, and the international collaborations that develop around shared data, are key to achieving the objectives of the CBD and the Nagoya Protocol. We remain concerned about any effort to constrain, regulate, or introduce delays in the long-standing scientific best practice of rapid, open exchange of information, including GSD. These concerns are particularly acute for situations where timely access is essential, such as identifying and managing invasive species. We urge Parties to be mindful of potential detrimental effects on public and private research and innovation, for biodiversity conservation and other sectors.

C. SUBMISSIONS FROM INTERGOVERNMENTAL AND OTHER RELEVANT ORGANIZATIONS AND STAKEHOLDERS

African Union Commission on behalf of the African Group
--

SUBMISSION

In a notification dated February 7, 2019, the CBD Secretariat, pursuant to decisions 14/20 and NP-3/12, respectively, invited Parties, other Governments, indigenous peoples and local communities, relevant organizations and stakeholders to submit views and information:

- (a) To clarify the concept, including relevant terminology and scope, of digital sequence information on genetic resources and if and how domestic measures on access and benefit-sharing consider digital sequence information on genetic resources;
- (b) On benefit-sharing arrangements from commercial and non-commercial use of digital sequence information on genetic resources.

In addition, Parties, other Governments and indigenous peoples and local communities were invited to submit information on “their capacity-building needs regarding the access, use, generation and analysis of digital sequence information on genetic resources, in particular for the three objectives of the Convention.

This paper represents the African Group’s response to these invitations.

(a) Views and information to clarify the concept, including relevant terminology and scope, of digital sequence information on genetic resources and if and how domestic measures on access and benefit-sharing consider digital sequence information on genetic resources.

I. Terminology

As noted by the CBD Draft Decision CBD/COP/14/L.36 in November 2018, there is an understanding that the phrase “digital sequence information” or “DSI” may not be the most appropriate phrase to capture the various types of information on genetic resources that may be relevant to the three objectives of the CBD/objective of the Nagoya Protocol. However, “DSI” is continuing to be used as a placeholder. There are, of course, other possible terms that could be used such as genetic information, natural information, genetic sequence data, and digital sequence data. In this regard we note with interest that on-going discussions to enhance the functioning of the multilateral system of ABS under the International Treaty on Plant Genetic Resources for Food and Agriculture have recently considered the possibility that DSI could be understood to be included in the terminology “associated available non-confidential descriptive information” used in Article 12.3.c) of the Treaty.

Our view is that a prolonged focus on terminology is not helpful for obtaining clarity on the concept of “DSI”. The goal of a focus on terminology is to narrow the scope of applicability of the Nagoya Protocol, and have certain subject matter excluded from it. Moreover, any definition on which agreement could be reached would likely be overtaken by rapid technological innovation and outdated in short order. In fact, the very nature of the controversy engaged in by the COP-MOP regarding the scope of the phrase “genetic resources” – despite the record of CBD negotiations clearly showing that the term was intended to include both material resources and the information they encode – exemplifies this eventuality.

While we favour the use of a neutral and wide term like “natural information” or “genetic information,” the African Group position is that the more relevant focus should be on utilization and not terminology, as precise terminology is not critical for this issue. This is because if information results from the utilization of genetic resources,⁵ it is within the scope of the CBD and Nagoya Protocol (in particular Article 5.1) and subject to benefit sharing, regardless of the terminology used to define the information. This is important because as technology continues to advance, new and unanticipated products of genetic resource utilization are likely to appear that might not be covered by terminology to be adopted in 2020.

If our negotiating partners nevertheless insist on further discussions around terminology we would suggest that a useful approach could be to deconstruct the placeholder term “DSI” into a continuum that starts with raw genetic sequence data obtained from primary scans of naturally occurring sequences and then progresses through compiled whole genomes to annotated or isolated functional genes, eventually culminating in useful discoveries and/or inventions that can be patented and/or used for gene editing or other forms of genetic manipulation. Such an exercise would be most useful if it were conducted in the context of negotiating differentiated benefit sharing rates for different classes of natural information utilisation.

II. Scope

There remain widely divergent views on whether “DSI” should be viewed narrowly or broadly. Our view is that assessing the scope of “DSI” involves a determination of its relationship to the phrases “genetic resources” and “benefits arising from the utilisation of genetic resources”. Several African countries

⁵ “Utilization of genetic resources” is defined in Article 2 of the Nagoya Protocol as “to conduct research and development on the genetic and/or biochemical composition of genetic resources, including through the application of biotechnology as defined in Article 2 of the Convention.”

currently include “DSI” in the definition of genetic material⁶ and thus of genetic resources contained in their national legislation, while many more countries have initiated changes to their ABS laws to ensure that DSI is clearly included. The following African countries’ approaches to DSI and ABS provide illustrative examples:

Ethiopia

The Ethiopian revised ABS proclamation emphasizes the role that communities play in the conservation, enhancement, development and sustainable use of biodiversity resources. It also underscores the need for fair and equitable sharing of benefits arising from the utilization of Ethiopian genetic resources.

The revised draft law elaborates the definition of ‘genetic resource’ as ‘any material of biological resource containing genetic information having actual or potential values for humanity and includes derivatives and digital sequence information’. The attending conditions of access to genetic resources have also been elaborated. Prior informed consent (PIC) of the institute (Ethiopian Biodiversity Institute) and the establishment of mutually agreed terms (MAT) is one condition of access. Likewise access to community knowledge is subject to the PIC of the concerned local community and the establishment of the MAT with the Institute.

Physical access to genetic resources by persons outside of the country (foreigners) in addition to the requirement of PIC and MAT is required to be accompanied by a personnel of the Institute or any other relevant body to be designated by the institute. Indeed this emphasizes the level of importance attached to access requirements by the governing law of Ethiopia. In cases of research the Ethiopian ABS law also requires that the research be based in Ethiopia with the participation of Ethiopian nationals unless it proves to be impossible to arrange for this. This condition is in relation to ensuring that non-monetary benefit sharing are also at hand by way of technology transfer in researches conducted. As a provider country Ethiopian law anticipates utmost collaboration between Ethiopia and the users of its genetic resources in an informed and transparent manner. Nonmonetary benefit sharing may among others also extend to joint ownership of intellectual property rights.

In the same token among the conditions of denial of access relate to instances where the requested access to the genetic resource is of an endangered species and where the access does not relate to research for rehabilitation of endangered species, the access is intended for use of genetic resources for purposes contrary to national laws of Ethiopia and where the genetic resource is of special cultural value or socioeconomic interest to the country.

In situations where the same genetic resource is found *in situ* within Ethiopian jurisdiction and other Parties transboundary cooperation is encouraged. Similar transboundary cooperation is required where traditional knowledge is shared by Ethiopia and other Parties to the Protocol. In such instances, considerations for common ABS frameworks with neighboring country Parties with the view of ensuring common equitable benefit sharing mechanisms based on common MAT is at hand.

Nevertheless, it is unclear how PIC would be obtained for sequence information from Ethiopian genetic resources (which may not be identified as such) which are already available in publicly and privately accessible databases hosted outside of Ethiopia. In addition, it is not evident how non-compliance with the PIC requirement for such sequence information would be tracked, or how benefit sharing obligations apply for both non-commercial and commercial uses of DSI.

Malawi

Malawi regulates access to, and benefit sharing of biological resources in accordance with the Environment Management Act (Cap. 60:02 of the Laws of Malawi) as well as other sectoral legislation. The Environmental Management Act states that *the biological and genetic resources of Malawi shall constitute an integral part of the natural wealth of the people of Malawi and shall be protected, conserved and managed for the benefit of the people of Malawi; and shall only be exploited or utilized in accordance with*

⁶ “Genetic material” is defined in Article 2 of the CBD as “any material of plant, animal, microbial or other origin containing functional units of heredity.”

the provisions of the Act or any other written law of Malawi. The Act empowers the Minister to facilitate development of legislative proposals, issue guidelines and prescribe measures for the protection, conservation and sustainable management and utilization of genetic resources; prescribe measures to regulate access to genetic resources by noncitizens or non-residents of Malawi; ensure PIC of communities is obtained and is an essential component for any arrangement in bio-prospecting, ensure effective equitable sharing of benefits and sustainable business mechanisms for the transfer of biotechnology; protect intellectual property rights of communities; prohibit or restrict any trade or traffic in any component; provide for fees payable in respect of accessing the resources and the export thereof; provide guidelines for reviewing of genetic materials and patenting requirements for indigenous species; regulate the collection, characterization, evaluation and documentation of plant genetic resources for food, agricultural and medicinal purposes; and regulate any other matters considered necessary for the sound management of the genetic resources.

Based on the provisions of the Act, Malawi has developed Guidelines on Access to Biological Resources and the Fair and Equitable Sharing of Benefits arising from their Use to provide clarity and certainty to Users and Providers of biological resources. The guidelines clearly indicate that Malawi considers all activities involving the collection and export and utilization of physical biological resources, traditional knowledge associated with genetic resources, use of genetic information or any forms of DNA/RNA sequences or sequence data in any format including in microbiological, digital or synthetic or in any other format associated with genetic resources to trigger benefit sharing obligations. Hence Users of DSI of genetic resources that originate from Malawi are required to comply with the provisions of the laws of Malawi.

Malawi has observed from most of the applications for access, export and utilization of biological resources that in most research projects there is an obligation to make sequence data or any other genetic information digitally available online. As such under current MATs a section addressing use of genetic information or any forms of DNA/RNA sequences or sequence data in any format is included.

ABS contracts indicate that in any digital publication of sequence data or any other digital expression of the genetic resources or results from the genetic resources, full acknowledgement is to be given to the Government of Malawi, the project enabled by the contract, and the following clause shall be enclosed in the digital publication: “*The government of Malawi has commercial rights or other further use rights in products or processes developed based on the research results or this DSI, and any use requires a contract of use with the Government of Malawi.*” Making research results or alike available online shall require the User of the information to be bound by a standard clause respecting the economic interests of Malawi, by a ‘click-wrap’, using a standard online accept-condition. Access shall be conditioned to the filling in of information about “name, affiliated institution, interest in accessing the digitalized information and the intended use” including an acceptance of the requirement in the Malawian Regulation of having a contract if embarking on using the information. The click-wrap shall send a cookies notifying Malawi about the one accessing the information.

DSI on genetic resources is also being included in Malawi’s ABS regulations which are under development because Malawi considers DSI and the use of genetic information or any forms of DNA/RNA sequences or sequence data in any format, including in microbiological, digital or synthetic or in any other format as a way of conveying, presenting or expressing genetic resources using a particular arrangement or sequences which still translate and provide useful information to the User on the functioning, characteristics and other traits of the original genetic resources.

Capacity building and technology transfer is one of Malawi’s priorities in benefit sharing arrangements. Malawi desires to develop capacity and technologies that allows it to participate in the value creation of biological resources and DSI. Malawi would like its researchers and communities to be co-inventors of products and processes from its genetic resources and relevant DSI and support its science and technology agenda. The impact of capacity building and technology transfer is based on the potential to provide long term benefits on transforming livelihoods and empowering communities. Benefit sharing funds have the potential to address some of the science and technology needs of Malawi through development of good research and development infrastructure and value creation centers for its genetic resources to generate

benefits and incentives for communities. The quest to obtain information on each of the thousands of genes, gene products and other characteristics of genetic resources indicates the demand for DSI for commercial and scientific users. Malawi needs capacity to develop its own data banks so that it is able to store and maintain this information in a way that will allow it to benefit from any value creation thereafter.

Conservation and ensuring sustainable use of biological resources requires financial, human and infrastructure investments. It is only fair that communities and countries that have dedicated their resources to ensure that genetic resources are available for the scientific and commercial users benefit from their contribution. Capacity is therefore also needed on how to help communities and institutions develop good business arrangements with Users of DSI for fair and equitable benefit sharing.

South Africa

The 2013 amendment to the National Environmental Management: Biodiversity Act, 2004 (Act No. 10 of 2004) ensures that the nation's indigenous genetic and biological resources are developed and utilized in an ecologically sustainable manner while promoting social and economic development in particular in the areas where the indigenous genetic or biological resources and associated traditional knowledge are accessed, as one of its intentions.

These amendments included the following definitions which are linked to DSI: “*Derivative*”, in relation to an animal, plant or other organism, means any part, tissue or extract of an animal, plant or other organism, whether fresh, preserved or processed, and includes any genetic material or chemical compound derived from such part, tissue or extract; “*Genetic material*”, means any material of animal, plant, microbial or other biological origin containing functional units of heredity; “*Genetic resource*”, includes any genetic material; or the genetic potential, characteristics or information of any species. These resources may be gathered from the wild or accessed from any other source.

PIC and establishment of the MAT are compulsory requirements for requesting commercial access to indigenous genetic and biological resources and also for access to traditional knowledge associated with indigenous genetic and biological resources.

South Africa believes that creation of sequence information from genetic resources depends on the physical access to the genetic resources. Therefore, the MAT and the permit templates contain mandatory clauses that address third party transfer terms and conditions which could include the utilization of DSI on genetic resources, whether stored in public or private databases. The majority of academic researchers are obliged to include sequence data in order for their research results to qualify for publication in peer-reviewed journals. Most journals require that the sequence data included in the paper are published in a globally accessible repository. However, there is no direct benefit-sharing for the use /access to such sequence data.

It is clear from the South African example that the Parties to the Nagoya Protocol on ABS should continue their work towards finding innovative/creative international policy solutions aimed at ensuring fair and equitable sharing of benefits with the country of origin of the original genetic resources which contributed to the creation of sequence data.

Uganda

In Uganda, the National Environment Act No. 5 of 2019 defines genetic resources to mean genetic material of actual or potential value. It is the intention of the government of Uganda to revise and update its National Environment (Access to Genetic Resources and Benefit Sharing) Regulations (mapped against the Bonn Guidelines) to further elaborate on the scope of the definition of genetic resources. But suffice it to say that the intention is to incorporate digital advancements in science and technology that are traceable to the original genetic resources accessed from Uganda.

So far, the discussions in this realm are that Uganda will not be held back by seeming uncertainly about how far the word “derivatives” can be stretched. The country is well aware that at this point in the discussion the elements of PIC and MAT are still very hazy; nevertheless, we are also equally aware that rapid advances in science and technology tends to outpace the development of regulatory tools. It is therefore

imperative to develop legislation that will enable the country to put in place measures to ensure benefit sharing.

The African Group consider DSI to result from the utilization of genetic resources, a view that is not limited to developing/provider countries.⁷ The African group is also of the view that the phrase “*genetic resources*” encompass DSI. Therefore, Uganda must consider what constitutes the genetic resources and work on a definition that explicitly makes reference to the genes or genetic material which constitute a genetic resource.

The above mentioned National Environment Act provides the legal basis for the sustainable management and utilization of the genetic resources of Uganda for the benefit of the people of Uganda. The said law requires the establishment of appropriate arrangements for access of genetic resources by non-citizens, measures for regulating the export and import of genetic resources and the sharing of benefits derived from genetic resources originating from Uganda.

The Act provides the basis for the revision and update of the National Environment (Access to Genetic Resources and Benefit Sharing) Regulations whose development was informed by the Bonn Guidelines on ABS. Now that Uganda has ratified the Nagoya Protocol and is a key negotiator on DSI, Uganda is convinced that it is imperative to develop national capacity to operationalize the Nagoya Protocol by way of ingraining the salient features of genetic resources and their derivatives in national legislation and in tracking use of Uganda’s genetic resources.

University training and capacity development of research institutions in this regard is important. It is equally important to develop the capacity of regulatory institutions for environment, science and technology, in terms of knowledge enhancement as well as provision of equipment. It is then that Uganda will adequately be able to understand the dimensions of her genetic resources and the multiple uses to which they are or can be employed. This is not only important for conservation and sustainable use purposes, but is also crucial for socio-economic development. Hence, it is important to understand the connection between understanding genetic resources and their derivatives and the other 2 objectives of the CBD, as well as their contribution to socio-economic development.

(b) Views and Information on benefit-sharing arrangements from commercial and non-commercial use of digital sequence information on genetic resources.

As explained above, the African Group is of the view that the phrase “genetic resources” or the phrase “utilization of genetic resources” should be deemed to encompass “DSI” and also making DSI subject to benefit sharing obligations. The African Group is aware of the many concrete benefits of the fairly open access to DSI that researchers around the globe enjoy currently. We also agree that everyone benefits from advancements enabled by open “DSI” sharing. However, changes in research patterns are expanding the use of DSI in place of tangible material, a decoupling which threatens to undermine the viability of the CBD/NP ABS scheme. In this regard it is important to consider the extent to which access to DSI is and can be used as a substitute for access to material genetic resources.

As such, we do not see such diffuse gains as sufficient to preclude monetary benefit sharing under the CBD/NP. In the same way that the public benefits from many patented technologies, yet the patent holders are still entitled to monetary benefits, providers of genetic resources and the “DSI” obtained from their utilization, are still entitled to the possibility of monetary and specific non-monetary benefits DSI utilization. It should be noted, however, that the African Group sees merit in exploring the development of a benefit sharing approach for DSI that would attach to commercialized products and not hinder academic research.

The African Group has noted with concern the misplaced focus of industry and academic users of genetic information on possible measures to control access to DSI. In our view it would be far more productive to

⁷ See, e.g., Kaspar Sollberger, *Digital Sequence Information and the Nagoya Protocol: Legal Expert Brief on behalf of the Swiss Federal Office for the Environment*, (Apr. 7, 2018) (concluding that “sequencing activities qualify as research and development, which makes them come under the concept of utilization of genetic resources according to Art. 2(c) of the Nagoya Protocol.”)

concentrate on ensuring benefit sharing when DSI is utilised. In the absence of a benefit sharing solution many African countries, as noted above, either already control access to DSI or have initiated measures to do so, as a way of ensuring that benefits arising from the utilisation of DSI are shared fairly and equitably.

(c) Information on capacity-building needs regarding the access, use, generation and analysis of digital sequence information on genetic resources

As indicated in the above country examples regarding domestic legislation, the African Group is of the view that practical capacity building needs such as the need for sequencing equipment and related training will be critical. Foundry support for synthesizing modified/synthetic sequences for R&D and economic development purposes will also be required to fill the capacity gaps regarding engagements with DSI. However, capacity building is just one sub-component of benefit sharing and cannot be a substitute for monetary benefit sharing when monetary benefits arise from utilization.

Country submissions will and should additionally stipulate further specific capacity-building needs with regards to transactions on DSI, taking into account national needs and priorities, and the capacity of any particular country to make productive use of capacity building and technology transfer.

Finally by way of conclusion, considering the rate at which advances in technology are happening, the fate of biological resources rests on how well technology advances contribute to the conservation and sustainable use of biodiversity through ensuring that communities that have contributed to the conservation and generation of knowledge on the genetic resources benefit from the value creation of genetic resources and are not left behind as technology continues to advance. As the discourse on DSI under the CBD continues, it is critically important to constantly be cognizant of the following:

1. As technology continues to advance, functional genomic studies continue to generate extraordinary amounts of information from genetic resources for scientific investigations and for commercial applications of biotechnology without benefitting the communities from whom the genetic resources have been obtained;
2. Sequence data and its by-products have been made easily accessible to Users but the benefits have not been made easily accessible to Providers; and
3. Biodiversity loss continues.

NOTE: Submitted by the African Group of Negotiators on Biodiversity-Ad Hoc Group on Digital Sequence Information.

Secretariat of the International Treaty on Plant Genetic Resources for Food and Agriculture (Plant Treaty)

SUBMISSION

In response to the request for views and information as contained in CBD COP Decision 14/20, paragraph 9, the Secretariat of the International Treaty on Plant Genetic Resources for Food and Agriculture (Plant Treaty) is pleased to provide updates on Plant Treaty processes regarding “digital sequence information (DSI)”.

There are four areas where which DSI is currently relevant or under consideration within the Plant Treaty processes, in relation to the specific items for which the CBD COP requested views and information, namely:

- The Multilateral System of Access and Benefit-Sharing (Multilateral System);
- The Global Information System of Article 17;
- Cooperation with the CBD;
- The Multi-Year Programme of Work of the Governing Body.

A) Terminology, scope, domestic measures

Cooperation with the CBD

In Resolution 9/2017 on cooperation with the CBD, the Governing Body of the Plant Treaty requested its Secretary to continue collaborating and, as appropriate, coordinate with the Secretariat of the CBD on issues related to DSI in order to promote coherence and mutual supportiveness in their respective activities. In this context, the Governing Body recognized that the term “digital sequence information” would be subject to further discussion. It also recognized that there would be a multiplicity of terms that had been used in this area (including, *inter alia*, “genetic sequence data”, “genetic sequence information”, “genetic information”, “dematerialized genetic resources”, “in silico utilization”, etc.) and that further consideration would be needed regarding the appropriate term or terms to be used.⁸

Enhancement of the Multilateral System

In 2013, the Governing Body of the Plant Treaty launched a process to enhance the functioning of the Treaty’s Multilateral System of Access and Benefit-sharing by establishing an *Ad Hoc* Open-Ended Working Group, which is tasked, *inter alia*, with developing measures for an increase in user-based payments and contributions to the Benefit-Sharing Fund in a sustainable and predictable long-term manner. The Working Group has been considering, among other matters, revisions to the Standard Material Transfer Agreement. The mandate of the Working Group is extended to the current biennium 2018-19. In the biennium, a meeting of the Working Group took place in October 2018 and the next meeting is forthcoming in June 2019.

At the meeting of October 2018, the Working Group considered the Co-Chairs’ proposed consolidated text for the revised Standard Material Transfer Agreement, as attached in Resolution 2/2017 of the Governing Body.⁹

The Working Group discussed whether to reflect issues related to DSI in the Draft Revised Standard Material Transfer Agreement. However, there are different views on the matter in the ongoing negotiations.

Multiyear Programme of Work (MYPoW)

In Resolution 13/2017 on the MYPoW, the Governing Body decided to consider, at its Eighth Session, the potential implications of the use of DSI for the objectives of the Plant Treaty, and to consider it for inclusion in the MYPoW at that Session. Based on the request by the Governing Body in the same Resolution, Contracting Parties, other governments, relevant stakeholders and individuals are providing information to the Governing Body, including on: terminology used in this area (*emphasis added*); actors involved with DSI on PGRFA; the types and extent of uses of DSI on PGRFA; the relevance of DSI on PGRFA for food security and nutrition.

The Secretariat of the International Treaty is publishing the submissions, as received, on its website.¹⁰ The submissions will also be compiled and made available to the Contracting Parties in preparation for the discussion of the issue at the next Eighth Session of the Governing Body.

B) Benefit-sharing arrangements from commercial and non-commercial use

As benefit-sharing, both monetary and non-monetary (including information), is a component of the Multilateral System, the on-going consideration of DSI in the context of the enhancement of the Multilateral System may be relevant to such forms of benefit-sharing.

As indicated above, the implications of the use of DSI on the objectives of the Plant Treaty, including the fair and equitable sharing of the benefits arising out of the use of PGRFA, will be considered by the Governing Body at its next Eighth Session in the context of the MYPoW.

Global Information System on Plant Genetic Resources for Food and Agriculture (GLIS)

⁸ Resolution 9/2017 is available at: <http://www.fao.org/3/a-mv088e.pdf>

⁹ Resolution 2/2017 is available at: <http://www.fao.org/3/a-mv104e.pdf>.

¹⁰ <http://www.fao.org/plant-treaty/overview/mypow/dsi/en/>.

Article 17 of the Plant Treaty establishes the GLIS to facilitate the exchange of PGRFA information, with the expectation that such exchange will contribute to the sharing of the benefits by making information available. The Vision and Programme of Work of the GLIS that the Governing Body endorsed in 2015, include the promotion of transparency of rights and obligations related to accessing, sharing and using information associated with germplasm, and the establishment of ways to exercise those rights and obligations within the GLIS. Under such component of the Vision and Programme of Work, the subject of access to, and use of, DSI/genetic sequence data (GSD) on PGRFA has been under consideration in the past and current biennia. In parallel with the consideration of DSI/GSD, the deployment of Digital Object Identifiers (DOIs) has been prioritized in order to facilitate the exchange of PGRFA information based on existing information systems.¹¹

In Resolution 5/2017 on the implementation of the GLIS, the Governing Body requested the Scientific Advisory Committee on the GLIS (SAC) to consider DSI, as far as DSI are generated from the use of PGRFA and related to the implementation of GLIS.¹² At its meeting of June 2018, the SAC requested the Secretariat of the Plant Treaty to gather information from users of the GLIS, including the Centers of the Consultative Group on Agricultural Research (CGIAR Centers) and other institutions managing crop germplasm repositories, on the current application of DOIs to crop germplasm in the Multilateral System for which DSI/GSD are available in compatible information systems.¹³ In most of the genbank information systems, only passport data and phenotypic data are available. Based on information gathered so far by the Secretariat, it appears that the use of DOI for plant germplasm accessions is increasingly accepted by genebanks. In the views of some stakeholders, the connection of these data with genetic information via DOIs, however, is very promising and a task for the Global Information System could be to promote the use of the accession DOIs as metadata for sequencing/genotyping consortia, thus improving the later integration with these data.

Consortium of European Taxonomic Facilities (CETAF)

SUBMISSION

Preamble

At the 2018 Conference of the Parties to the CBD and Meeting of the Parties to the Nagoya Protocol in Egypt decisions were adopted to gather information on Digital Sequence Information (DSI) in order to inform further discussion by the Parties. The decision requests Parties, other Governments, indigenous peoples and local communities, relevant stakeholders and organizations to submit their views and information:

- (a) To clarify the concept, including relevant terminology and scope, of digital sequence information on genetic resources and if and how domestic measures on access and benefit-sharing consider digital sequence information on genetic resources;
- (b) On benefit-sharing arrangements from commercial and non-commercial use of digital sequence information on genetic resources.

The areas of research that use “DSI” are characterised by rapid growth and innovation. The research carried out by CETAF members in the realms of taxonomy, systematics, description, environmental and biodiversity studies, evolutionary development, conservation support etc. are no different in this respect. Terminology applied to “DSI” modalities of its use when agreed need to be unambiguous and ‘future-

¹¹ More information on DOIs can be found at: <http://www.fao.org/plant-treaty/areas-of-work/global-informationsystem/doi/it/>.

¹² Resolution 5/2017 is available at: <http://www.fao.org/3/a-mv103e.pdf>.

¹³ IT/SAC-GLIS-3/18/Report, paragraph 22, available at: <http://www.fao.org/3/CA0526EN/ca0526en.pdf>. In the report of the SAC, the term “digital sequence information” was utilized interchangeably with “genetic sequence data”, without any prejudice to the possible definition of terminology by the Governing Body.

proofed' to whatever extent possible. This will provide certainty and a firm base to carry on research and benefit sharing.

This consideration will affect the characterisation of “DSI”, and results in our proposal below.

Currently, the vast majority of research results from all countries are made publicly available, as are the data used in scientific studies, which includes “DSI”. The ability to replicate other researchers' results is a cornerstone of all scientific research, and this requires the data to be available. That “DSI” is freely available has raised concerns among some developing countries that it could be used in commercial applications without triggering an obligation to share benefits with the provider countries from which the original genetic resources were acquired. CETAF members understand this concern (although this does not apply to the vast majority of research carried out by CETAF members, which is non-commercial), but also wish to highlight the large number of non-monetary benefits which are shared by the scientific community which actively contribute to and support the objectives of the CBD. These benefits were discussed in our submission to the Executive Secretary in 2017¹⁴ and will not be repeated here.

Research data including “DSI”, when published, are maintained to the standardised quality norms of the global research community and available for use in Provider and User Countries at zero marginal cost. This important function of science is reflected in Article 15 of the CBD, which calls Parties to take legislative, administrative or policy measures ... with the *aim of sharing in a fair and equitable way the results of research and development and the benefits arising from the commercial and other utilization of genetic resources with the Contracting Party providing such resources*.

CETAF is committed to benefit-sharing and has developed best practices for its members to facilitate this. For “DSI”, the ABS Core Group believe that the most effective basis for benefit-sharing is on a global basis for the common good, in the manner required by Aichi Target 19. This will require a functional common set of technical standards. Promising practical examples demonstrating such standards are available, including the International Nucleotide Sequence Data Collaboration (INSDC), Global Biodiversity Information Facility (GBIF) and the Barcode of Life Database (BOLD), which address both technical and legal issues, in the case of GBIF explicitly operating within an intellectual property rights framework. We think building on these should be preferred instead of implementing new systems that aim to restrict and regulate DSI. Because of the exponentially growing amount of data, developing new systems with additional legal and policy-related requirements would be a difficult and expensive task with unknown results but potentially negative impact on science globally, and particularly on CBD implementation.

This submission builds on the 2017 submission from CETAF to the Executive Secretary on Digital Sequence Information on genetic resources – benefits of their use and their public availability for the three objectives of the Convention on Biological Diversity, and ramifications of restricting access to DSI¹⁵.

Contents

Comments in Detail

1. The concept, including relevant terminology and scope, of “digital sequence information” on genetic resources
2. Domestic measures on access and benefit-sharing considering “digital sequence information” on genetic resources;
3. Benefit-sharing arrangements from non-commercial use of “digital sequence information” on genetic resources.

Summary

Comments in detail

- 1. The concept, including relevant terminology and scope, of “digital sequence information” on genetic resources**

¹⁴ <https://www.cbd.int/abs/DSI-views/CETAF-DSI.pdf>.

¹⁵ <https://www.cbd.int/abs/DSI-views/CETAF-DSI.pdf>

The term Digital Sequence Information is not used by CETAF members in their work and appears to be limited to policy discussions¹⁶. The very diverging interpretations of the term now current and the resultant ambiguity make it important to use a different term of precise meaning.

One issue lies in the term ‘information’. We understand that while ‘data’ are observations of naturally occurring states lacking extrapolated meaning ‘information’ emerges through cognitive (or other) processing and application of data. In the “DSI” context ‘data’ are extracted from naturally-occurring genetic resources – the genetic resource itself when accessed has no intrinsic ‘information’.

The relevant data are the arrangement of nucleotides on strands of naturally occurring DNA or RNA¹⁷. This is ‘Nucleotide Sequence Data’ (NSD). Information about the genetic resource arises through analysis of these data.

Free sharing of both data and information, so they are available to researchers and other users in Providing and other Countries, is the usual practice in the non-commercial sector, and we view such provision as constituting a shared benefit. However, there are important differences between the two.

In practice, since ‘information’, is developed through analysis of ‘data’, it is potentially covered by Intellectual Property Rights (IPR). Sharing of IPR is identified as an example of a non-monetary benefit in the Nagoya Protocol Annex; consequently, it can be seen that IPR are a result of research and not under sovereign rights of a country as are natural resources (genetic resources).

Analysis of NSD is a key tool for taxonomy. The data used for analysis are aggregated from naturally occurring GR and downloaded from databases such as GenBank and the other INSDC¹⁸ members’ databases, and include non-coding and coding sequences, regulatory sequences, conserved sequences, genes that encode specific traits, and ‘junk’ DNA¹⁹; sequences do not need to have a known function to be of relevance in our research. There is no maximum size for a usable sequence. Analysis might be of single genes, multiple genes, entire genomes of organisms, of a clade (pangenome) or environmental samples (metagenomes). The results of analysis are interpreted to further our understanding of biological diversity.

We propose that discussion on “DSI” distinguishes between data (NSD), and information. We also propose that the concept of DSI be restricted to NSD. We note that this clear concept is also in line with ‘Genetic Sequence Data’ (GSD) as proposed by some Parties to the CBD. We specifically exclude ‘digital’ from our proposed terminology to avoid inappropriate restriction to a single current means of data storage and transmission of aggregated data from GR.

For the vast majority of taxonomic and systematic research, it is not the discovery and application of functions of the genes per se that are important. In any case, since the function can only be discovered by experiment or through the application of existing knowledge to predict or test for function (i.e. extrapolation of data including through automatic means) this would come under the heading of information.

Sequence data may be associated with a set of other data to increase its scientific value, such as:

- i) Collection site of the organism or sample from which the NSD was obtained;
- ii) the date on which it was collected;
- iii) the name of the collector;
- iv) the place where a physical voucher is stored (if it is retained) and the unique identifier of that voucher;

¹⁶ Laird & Wynberg, 2018, A Fact-Finding and Scoping Study on Digital Sequence Information on Genetic Resources in the Context of the Convention on Biological Diversity and the Nagoya Protocol. CBD/DSI/AHTEG/2018/1/3

¹⁷ Nucleotides are the subunits that are connected into long chains to make nucleic acids (DNA and RNA). The four types of nucleotides in DNA are Adenine, Thymine, Guanine, and Cytosine, and in RNA Thymine is replaced by Uracil. The five nucleotides are usually abbreviated to A, T, G, C and U. The order in which these nucleotides occur in a strand of DNA or RNA is the DNA or RNA sequence or Nucleotide Sequence.

¹⁸ International Nucleotide Sequence Database Consortium.

¹⁹ Sequences currently of no known function

- v) the taxonomic name of the organism from which the DNA was sequenced.

While we do not consider this to be NSD, it is helpful contextual information and, where appropriate (and where it exists) it can be made available with NSD to which it applies. Associating these data is scientific best practice, but far from all sequences stored in public databases are associated with all of these data. Permit conditions may be stored as part of the record.

2. Domestic measures on access and benefit-sharing considering digital sequence information on genetic resources

We are aware that some countries are introducing domestic legislation including NSD. This gives us cause for concern. One is terminology; countries are using a variety of terms that are sometimes of unclear meaning, are often used inconsistently and thus increase legal uncertainty, instead of reducing it. Another is practical. CETAF members expect to seek permits or their equivalent when collecting biological specimens for research or addition to collections, and to share agreed benefits. However, these are negotiated and set out in bilateral arrangements, made in the context of significant investment of time and resources to establish scientific work in one country. Use of NSD includes work that is very different. Often the research is not targeted at a particular country but at a particular species or group of species (perhaps hundreds or even thousands of species). The uses to which the NSD are put include for example simple comparison to identify an organism or taxon from which a similar sequence has been isolated (or such a sequence within an eDNA sample, for example), developing hypotheses of evolutionary relationships and assessment of biodiversity richness. In the overwhelming majority of cases in our sector the outputs are information-based and shared globally. If countries require bilateral agreements before NSD are used this is likely to have several detrimental outcomes. Some researchers may simply ignore the requirement, exacerbating mistrust and leading to breakdown of both relationships and the growing incorporation of ABS in research workflows globally. Others will develop algorithms to avoid use of NSD from countries making such requirements, leading to fewer data and less information being generated to address biodiversity management priorities. Such a limitation would compromise achievement of the Aichi Targets as well as national NBSAP goals. The reasons for these actions is that the time required to reach an ABS agreement is likely to damage research significantly; such delays for access to physical specimens are already problematic, and there is no reason to expect that the very much higher rate of use of NSD will be handled any more expeditiously. To put this into context, every weekday, well over 38 million requests are made to EMBL-EBI websites²⁰.

3. Benefit-sharing arrangements from non-commercial use of digital sequence information on genetic resources.

In our submission of 2017 we discussed the value of DSI to the implementation of the CBD, and would withdraw nothing from that statement²¹, which was based almost entirely on the concept of NSD discussed here. Implementation of the CBD is of course on a national basis, and thus countries implementing the CBD are making use of benefits developed through the use and generation of NSD.

CETAF members agree bilateral benefit-sharing arrangements as a normal part of Mutually Agreed Terms when accessing genetic resources. Insofar as the benefits are data or information (as opposed to capacity-building, for example), practice is to both share them bilaterally and to publish them so they are available globally. While we have many examples of bilateral benefit-sharing in the context of tangible genetic resources, we have no experience of benefit-sharing on a bilateral contractual basis exclusively for NSD.

Although contract-based bilateral benefit-sharing for NSD is not in our experience, global benefit sharing from analysis of NSD is a common experience and part of daily work. As noted above, the prevailing model of non-commercial scientific work is of open sharing of data and outputs. As a concomitant to publication of research results researchers are required to make NSD from their own databases publicly available once research is published, generally through the large public databases of the INSDC. Thus benefit sharing when these benefits are new genetic sequences is done on a global (multilateral) basis rather than a bilateral

²⁰ <https://www.ebi.ac.uk/about/our-impact>

²¹ <https://www.cbd.int/abs/DSI-views/CETAF-DSI.pdf>

basis. This methodology is far more efficient and valuable to all users, since it allows access to sequences relating to species outside national borders, important for identification of invasive species, for example²².

The global resource of NSD provided by the INSDC is in constant use worldwide; although the majority of users come from North America, Europe and China the data made available are used by every Party to the CBD²³ (see <https://www.ebi.ac.uk/about/our-impact> for a real-time visualisation of use of EMBL databases). While researchers can (and sometimes do) send direct links to uploaded NSD to provider countries, there is no evidence that this makes any difference to availability and use of the data developed and shared.

NSD are in use globally, but there are still capacity-building needs to increase Parties' abilities to realise the benefits and exploit these data. The SCBD has supported training in DNA barcoding, which includes making use of the NSD in the BOLD system. INSDC members also offer training and a range of training materials. CETAF member organisations are also active in capacity building. This may take the form of training as a part of research, for example training students while working in labs in providing countries, joint research involving generation and analysis of NSD, in-house training at bachelor's, master's, and PhD levels, and informally through professional contact. Many CETAF members also run DNA labs as a part of their infrastructure, and make these available to visitors and colleagues from developing countries, effectively increasing the capacity of those countries.

Summary

The term "Digital Sequence Information" is ambiguous and "DSI" is increasingly used as a convenient acronym without consideration of what it encompasses. We suggest that a replacement term be used in discussion and negotiation – 'Nucleotide Sequence Data' (NSD). This is the order in which nucleotides (Adenine, Thymine or Uracil, Guanine, and Cytosine) occur in a strand of DNA or RNA. The definition excludes the term 'information' which is developed through analysis of the data and which might be under the Intellectual Property Rights of the researcher. It also excludes 'Digital' to avoid restriction to a single storage medium.

Research increasingly involves generation of new NSD and downloaded NSD from public databases. While we have no examples of bilateral benefit-sharing on a contractual basis for use of NSD, the prevailing model of scientific publication of research results and the underlying data means that these results, and NSD, are available globally, the NSD being open access. We have seen evidence that users in all countries are accessing these data. This global availability of information to assist countries in implementing the Convention on Biological Diversity has been called for in a number of COP decisions and under Aichi Target 19.

We are aware that countries do not all have sufficient capacity to make full use of NSD. We regularly engage in capacity building through training and joint research, and see this as a continuing activity.

We are concerned that the development of restrictions on use of NSD will damage biodiversity research. Biodiversity loss is alarming²⁴, and further restrictions to identify and understand Biodiversity will generate massive drawbacks for the people's well-being and all life on Earth and consequently endanger research in the conservation and sustainable use of Biodiversity.

NOTE: CETAF ABS Core Group

²² Lack of such sequences has been identified as a problem for invasive Alien Species detection – see Lyal & Miller, 2018, Capacity of United States federal government and its partners to rapidly and accurately report the identity (taxonomy) of non-native organisms intercepted in early detection programs. 22pp. https://www.doi.gov/sites/doi.gov/files/uploads/lyal_federal_capacity_taxonomy_contractorreport_22october2018.pdf

²³ Leibnitz Association, 2018, The DSI debate: a primer on the science and infrastructure behind DSI. Discussion paper.

²⁴ <https://www.ipbes.net/news/media-release-biodiversity-nature%E2%80%99s-contributions-continue-%C2%A0dangerous-decline-scientists-warn>

Consortium of German Natural History Collections (Deutsche Naturwissenschaftliche Forschungssammlungen), German Life Sciences Association (Verband Biowissenschaften, Biologie und Biomedizin in Deutschland) and the Leibniz Biodiversity Research Alliance (Leibniz Verbund Biodiversität)

SUBMISSION

Summary

The term “Digital Sequence Information” is ambiguous and “DSI” is increasingly used as a convenient acronym stemming from policy discussions without a clear concept of what it encompasses and is a term simply not used by scientists. We suggest that a replacement term be used in discussions and negotiations – ‘Nucleotide Sequence Data’ (NSD). This is the order in which nucleotides (Adenine, Thymine or Uracil, Guanine, and Cytosine) occur in a strand of DNA or RNA. The definition excludes ‘information’ which is developed through analysis of the data and which might be under the Intellectual Property Rights of the researcher. It also excludes ‘Digital’ to avoid restriction to a single storage medium.

Research increasingly involves generation of new NSD and heavily relies on NSD downloaded from public databases. The prevailing model of scientific publication of research results and the underlying data means that these results, and NSD, are available globally, the NSD being open access. Users in all countries access and use these data. This global availability of information to assist countries in implementing the Convention on Biological Diversity has been called for in a number of COP decisions and under Aichi Target 19.

We are aware that some countries do not have sufficient capacity to make full use of NSD. We regularly engage in capacity building through training and joint research, and see this as a continuing activity.

We are concerned that the development of restrictions on use of NSD will damage biodiversity research. Biodiversity loss is alarming* and further restrictions to identify and understand biodiversity will generate massive drawbacks for the well-being of mankind and all life on Earth.

About

The research carried out by scientists represented through Consortium of German Natural History Collections, DNFS (Deutsche Naturwissenschaftliche Forschungssammlungen), German Life Sciences Association (Verband Biowissenschaften, Biologie und Biomedizin in Deutschland, VBIO e. V.) and the Leibniz Biodiversity Research Alliance (Leibniz Verbund Biodiversität, LVB) focuses on biodiversity-related topics that directly or indirectly support the knowledge necessary to protection and sustainable use of biodiversity.

This joint submission is based on earlier views submitted to the CBD Executive Secretary on Digital Sequence Information on genetic resources by VBIO²⁵ 2017, the Leibniz Association²⁶ and the submission from CETAF²⁷ which was written with support of members of DNFS-institutions. We believe that both the benefits arising from the use of “digital sequence information” (“DSI”) and free, open access to “DSI” are vital for the three objectives of the Convention on Biological Diversity (CBD), and caution that restricting access to “DSI” in any way would have negative ramifications.

1. The concept, including relevant terminology and scope, of “digital sequence information” on genetic resources

→ **Replace “DSI” with Nucleotide Sequence Data (NSD)**

²⁵ <https://www.cbd.int/abs/DSI-views/VBIO-DSI.pdf>

²⁶ <https://www.cbd.int/abs/DSI-peer/Leibniz.pdf>

²⁷ <https://www.cbd.int/abs/DSI-views/CETAF-DSI.pdf>

The concept, including relevant terminology and scope, of ‘digital sequence information’ on genetic resource as a technical term seems to be limited to policy discussions²⁸ but is not used by scientists. This leads to very divergent interpretations of “DSI” in the current debate and huge ambiguities. We thus analyse the potential meaning “digital sequence information” in a scientific context first, and suggest the usage of a different term of precise meaning.

It is important to distinguish between ‘information’ and ‘data’. While ‘data’ are observations of naturally occurring states lacking extrapolated meaning, ‘information’ arises out of processing and application of data through cognitive efforts. The genetic resource itself when accessed has no intrinsic ‘information’, but in the “DSI” context, contains ‘data’ that are extracted from naturally occurring genetic resources, i.e. the arrangement of nucleotides on strands of naturally occurring DNA or RNA²⁹. This is ‘Nucleotide Sequence Data’ (NSD). ‘Information’ about the genetic resource arises through the subsequent research with NSD, and huge amounts of this emerging ‘information’ have relevance and importance for reaching the goals of the CBD and contributes to non-monetary benefit sharing already. For example, such studies can be used to support species conservation, enable more rapid biodiversity assessments, and to develop hypotheses of evolutionary relationships and assessment of biodiversity richness

.Furthermore, analysis of NSD is a fundamental requirement for basic research. The data used for analysis are aggregated from naturally occurring GR and downloaded from INSDC databases and include non-coding and coding sequences, regulatory sequences, conserved sequences, genes that encode specific traits, and ‘junk’ DNA³⁰ (the pure arrangement of nucleotides that does not have a known function). There is no maximum size for a usable sequence. Analysis might be of single genes, multiple genes, entire genomes of organisms, of a clade (pangenome) or environmental samples (metagenomes). The results of analysis are interpreted to further our understanding of biological diversity.

We recommend that discussion on “DSI” distinguishes between data (NSD), and user-generated information (which requires significant up-front investments before potential benefits of any kind can be generated). We also propose that the concept of “DSI” be explicitly and exclusively linked to NSD. We note that this clear concept is also in line with ‘Genetic Sequence Data’ (GSD) as proposed by some Parties to the CBD, however, referring to the cautioning remarks of the official Canadian submission (29 May 2019) on “genetics” and “genomics” and clearly prefer NSD instead of GSD, which seems to us the more precise term. We specifically exclude ‘digital’ from our proposed terminology to avoid inappropriate restriction to a single current means of data storage and transmission of aggregated data from GR. For the vast majority of scientific research, it is not the discovery and application of functions of the genes per se that are important. In any case, since the function can only be discovered by experiment or through the application of existing knowledge to predict or test for function (i.e. extrapolation of data including through automatic means) this would come under the heading of information.

Sequence data may be associated with a set of other data to increase its scientific value, such as:

- i) Collection site of the organism or sample from which the NSD was obtained;
- ii) the date on which it was collected;
- iii) the name of the collector;
- iv) the place where a physical voucher is stored (if it is retained) and the unique identifier of that voucher;
- v) the taxonomic name of the organism from which the DNA was sequenced.

²⁸ Laird & Wynberg, 2018, A Fact-Finding and Scoping Study on Digital Sequence Information on Genetic Resources in the Context of the Convention on Biological Diversity and the Nagoya Protocol. CBD/DSI/AHTEG/2018/1/3

²⁹ Nucleotides are the subunits that are connected into long chains to make nucleic acids (DNA and RNA). The four types of nucleotides in DNA are Adenine, Thymine, Guanine, and Cytosine, and in RNA Thymine is replaced by Uracil. The five nucleotides are usually abbreviated to A, T, G, C and U. The order in which these nucleotides occur in a strand of DNA or RNA is the DNA or RNA sequence or Nucleotide Sequence.

³⁰ Sequences currently of no known function

While we do not consider this to be NSD, contextual information is helpful and, where appropriate (and where it exists) it can be made available with NSD to which it applies. Associating these data is scientific best practice, but far from all sequences stored in public databases are associated with all of these data. Permit conditions may be stored as part of the record only if the permits are available as an IRCC through the ABS-Clearinghouse where a DOI can be generated. PDF permits (PICs/MATs) are at present not directly linkable to sequence data through the INSDC databases.

2. Domestic measures on access and benefit-sharing considering digital sequence information on genetic resources

We have worked closely with host countries to determine what may be sequenced during the course of our research projects and have placed importance on the fact that data must be published in open access databases for the broader benefit of science, and especially, so that our by partner scientists in-country can cite and re-use this important data and build their careers using this data. This is a clear win-win for all involved because it is only through shared learning about biodiversity that we will be able to achieve the targets agreed to by all. It is worth mentioning though that in the early stages of research, it is often not yet clear which particular species or group of species will be identified or isolated and, as such, it is unclear which sequences are potentially covered by a PIC/MAT/permit. For example, when sequencing mixed environmental DNA samples from microorganisms there is no *a priori* information about which organisms could be there.

We also warn against an overly protective stance by countries that strive to regulate NSD as we have observed that this, perhaps counter intuitively, ultimately leads to fewer data and less information being generated to address biodiversity management priorities. Such a limitation would compromise achievement of the Aichi Targets as well as national Biodiversity Strategies and Action Plans goals.

3. Benefit-sharing arrangements from non-commercial use of digital sequence information on genetic resources.

The main point of distinction between “DSI” and NSD is that ‘information’ is developed through analysis of ‘data’ and potentially is covered by Intellectual Property Rights (IPR). Consequently, it can be seen that IPR are a result of research and not under sovereign rights of a country as are natural resources (genetic resources). The free sharing of these sequence-based analytical results without claiming IPR is identified as an example of a non-monetary benefit in the Nagoya Protocol Annex and the scientific backbone for reaching the first two goals of the CBD, meet the Aichi Targets, and presumably enable the upcoming post-2020 Biodiversity Framework. In our submissions in 2017^{31,32}, we emphasized the value of “DSI” to the implementation of the CBD, and we continue to stand behind these statements.

Benefit sharing from analysis of NSD is fundamental principle of basic research and a common feature of the daily work of scientists globally. Open sharing of data and outputs is the prevailing model of non-commercial scientific work. The basic principles of good scientific practice require that data be made freely available to the scientific community so that the results can be replicated and validated. In the case of “DSI”, this is done by uploading sequence data to large sequence databases such as INSDC that guarantee free (to the user), unrestricted, worldwide availability, often known as “open access”. Databases such as the INSDC are used by scientists from Provider and User Countries and are maintained by the hosting countries (US, EU, Japan), thus offering both a monetary and nonmonetary especially for Providing Countries. Benefits arising from the use of NSD are usually shared as soon as they arise, i.e. when they are published. This methodology is far more efficient and valuable to all users, since it allows access to sequences relating to species outside national borders, important for identification of invasive species³³. We concur that benefit

³¹ <https://www.cbd.int/abs/DSI-views/VBIO-DSI.pdf>

³² <https://www.cbd.int/abs/DSI-peer/Leibniz.pdf>

³³ Lack of such sequences has been identified as a problem for invasive Alien Species detection – see Lyal & Miller, 2018, Capacity of United States federal government and its partners to rapidly and accurately report the identity (taxonomy) of non-native organisms intercepted in early detection programs. 22pp.

https://www.doi.gov/sites/doi.gov/files/uploads/lyal_federal_capacity_taxonomy_contractorreport_22october2018.pdf

sharing arrangements as a normal part of Mutually Agreed Terms when accessing genetic resources can and should be bilateral. A two-tiered system for NSD would create a yet more administrative and bureaucratic burdens that we fear would lead to a near paralysis of international research collaborations.

Every country in the world has scientists that use freely accessible NSD via platforms such as INSDC actively. Usage of these websites is global and is accessed and used literally by every country in the world³⁴³⁵. This contradicts the argument that “DSI” from Provider Countries is being exploited by user countries. Furthermore, the vast majority of NSD is created from human resources and GR that has its origin in the Global North. In order to have greater participation in the origin and usage of NSD by the Global South, access to NSD should be free and open for researchers around the world. NSD are used globally, but there are still capacity building needs to increase Parties’ ability to realise the benefits and exploit these data. Although the policy and technical details are challenging, this model of Open Access has the enormous advantage that the societal challenges already mentioned above and the first two CBD goals will continue to be addressed and the international scientific community can continue to work together. In order to increase capacity building in NSD, the capacity building should be intensified. The SCBD has supported training in DNA barcoding, which includes making use of the NSD in the BOLD system. MOOC (massively open online course) could be coordinated with the INSDC databases and/or new sequencing centres could, and the existing training of INSDC members and a range of training materials could be expanded. DNFS, VBIO and LVB member organisations are also active in capacity building. This may take the form of training as a part of research, for example training students while working in labs in providing countries, joint research involving generation and analysis of NSD, in-house training at bachelor’s, master’s and PhD levels, and informally through professional contact. Many institutions represented by DNFS, VBIO and LVB also run DNA labs as a part of their infrastructure, and make these available to visitors and colleagues from developing countries, effectively increasing the capacity of those countries.

4. *Further Remarks*

- **“DSI” is important for achieving CBD related goals**

Research data including “DSI”, when published, are maintained to the standardised quality norms of the global research community and available for use in Provider and User Countries at zero marginal cost. Free, unrestricted access to such data is essential not only for the achievement of the first two objectives of the CBD, the Aichi Target****s and the post-2020 Biodiversity Framework††††, but also for human, animal, and plant health especially during new outbreaks as it enables short and long-term analyses including epidemiology, diagnosis, and monitoring.

- **Life sciences depend on unrestricted access to sequence information**

Publication of research results from all countries including the data used in scientific studies and molecular data aggregated through utilisation of genetic resources, is required by peers and journals alike, in order to verify or replicate research results. Life science globally depends on regular and unrestricted access to sequence information through large public databases. The fact, that such data is freely available raised concerns among some developing countries that “DSI” could lead to commercial applications without triggering obligation to share benefits with the provider country. Even though the vast majority of the research carried out by our members is of non-commercial nature, we understand this concern. Nevertheless we want to emphasise the *huge* amount of non-monetary benefits which our scientific community actively contributes to and supports the objectives of the CBD. Furthermore, these benefits are directly shared with international partners and collaborating scientists that, without the unrestricted use of this data, would be excluded from current research and unable to access the data that we jointly produce and publish. These benefits were discussed in the submission of the Consortium of the European Taxonomic Facilities (CETAF) to the Executive Secretary of the CBD in 2017³⁶, which we fully endorse. Research data including

³⁴ Leibnitz Association, 2018, The DSI debate: a primer on the science and infrastructure behind DSI. Discussion paper.

³⁵ see <https://www.ebi.ac.uk/about/our-impact> for a real-time visualisation of use of EMBL databases)

³⁶ <https://www.cbd.int/abs/DSI-views/CETAF-DSI.pdf>

“DSI”, when published, are maintained to the standardized quality norms of the global research community and available for use in Provider and User Countries at zero cost to the users. Free, unrestricted access to such data is essential not only for the achievement of the first two objectives of the CBD, the Aichi Target§§§§§§ and the post2020 Biodiversity Framework*****, but also for human, animal, and plant health especially during new outbreaks as it enables short and long-term analyses including epidemiology, diagnosis, and monitoring. Article 15 of the CBD reflects this important function of science and calls Parties to take “*legislative, administrative or policy measures (...) with the aim of sharing in a fair and equitable way the results of research and development and the benefits arising from the commercial and other utilization of genetic resources with the Contracting Party providing such resources*”.

- **“DSI” as a common good**

It is our opinion, that the most effective basis for benefit-sharing on a global scale is “DSI”, as a basis for the common good, in the manner required by Aichi Target 19. We need a functional common set of technical standards to achieve this, and promising practical examples demonstrating such standards are already available include the International Nucleotide Sequence Data Collaboration (INSDC), the Global Biodiversity Information Facility (GBIF) and the Barcode of Life Database (BOLD). All these address both technical and legal issues and in the case of GBIF explicitly operating within an intellectual property rights framework. Instead of developing and implementing new systems to restrict and regulate “DSI” with unknown outcomes and high risk of failure, we believe that building on established principles should be preferred. In this context it is relevant to understand that operation and maintenance of such public databases storing “DSI” is a huge task (i.e., billions of US dollars over several decades), and that data uploaded to INSDC††††† are mirrored for example among GenBank and the other INSDC members’ databases on different servers in several countries around the globe on a daily basis. Thus, the same datasets are stored and exchanged simultaneously on servers in multiple countries, which will cause additional technical difficulties in regulation. Because of the amount of the existing and exponentially growing quantity of data, developing new systems with additional legal and policy-related requirements would be a difficult and expensive task with unknown results but potentially negative impact on science globally, and particularly on CBD implementation.

- **Avoid ambiguity and legal uncertainty**

Any terminology that resulting from the “DSI” discussion as well as the modalities of the use of terms has to avoid ambiguities and need to be ‘future-proofed’ to whatever extent possible – both in terms of administrative burden and impeding scientists in provider countries to participate in the global community where open access to data is a prerequisite to publish and participate. Both are essential to ensure certainty and a firm base for research and benefit sharing and thus characterisation of DSI in the focus of our submission.

Chartered Institute of Patent Attorneys (CIPA)

SUBMISSION

Digital Sequence Information on Genetic Resources: Submission of views

We provide a submission of views pursuant to Decision 14/20, paragraph 9, of the fourteenth meeting of the Conference of the Parties to the Convention on Biological Diversity and Decision NP-3/12 of the third meeting of the Conference of the Parties serving as the meeting of the Parties to the Nagoya Protocol on Access and Benefit-Sharing.

About the Chartered Institute of Patent Attorneys (CIPA)

CIPA was founded in in 1882 and was incorporated in the United Kingdom by Royal Charter in 1891. It represents virtually all the 2000 or so registered patent attorneys in the UK, whether employed in industry or serving the general public. Total membership is over 3,200 and includes trainee patent attorneys, and other professionals with an interest in intellectual property (patents, trade marks, designs and copyright).

CIPA Members advise clients on a wide range of intellectual property matters, representing all types of enterprise both large and small in drafting, filing, prosecuting and enforcing patent rights throughout the world. Members are well placed to advise inventors who use genetic resources of their responsibilities under the Nagoya Protocol.

The Convention on Biological Diversity and the Nagoya Protocol

The Convention on Biological Diversity (the “CBD”) established three objectives: the conservation of biodiversity, the sustainable use of its components and the fair and equitable sharing of the benefits arising out of the utilization of genetic resources (Art. 1 CBD).

The Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization to the CBD (the “NP”) implemented and further specified Article 15 of the CBD, on access to genetic resources. The NP set up international rules governing access to genetic resources and associated traditional knowledge, and benefit sharing as well as user compliance measures. Benefit-sharing under the NP is based on so-called mutually agreed terms (MAT), which are contractual agreements concluded between a provider of genetic resources or traditional knowledge associated with genetic resources, and a natural or legal person accessing the genetic resource and/or associated traditional knowledge for the utilisation thereof (a “user”).

The key terms used in the CBD and, particularly in Article 2 of the NP, are as follows:

- “**Genetic resources**” means genetic material of actual or potential value.
- “**Utilisation of genetic resources**” means to conduct research and development on the genetic and/or biochemical composition of genetic resources, including through the application of biotechnology as defined in Article 2 of the CBD.
- “**Biotechnology**” means any technological application that uses biological systems, living organisms, or derivatives thereof, to make or modify products or processes for specific use.
- “**Derivative**” means a naturally occurring biochemical compound resulting from the genetic expression or metabolism of biological or genetic resources, even if it does not contain functional units of heredity.

DSI does not fall within the definition of ‘genetic resources’ as defined in the CBD and NP

The terminology used to describe DSI in the context of the CBD and NP is always in relation to information on a genetic resource. Hence, by definition DSI on genetic resources cannot be considered to be comprised within the genetic resource itself. By analogy, the information about a chemical element as presented in the periodic table is not considered to constitute a part of the element itself. Rather, information on a chemical element is simply a description of the properties that the element exhibits in nature. This is no different to DSI which, in its broadest form, may simply represent a description of properties of a physical material which happens to be a genetic resource. Clearly it is wrong to suggest that a digitised sequence of letters that are biology shorthand for a nucleic acid sequence of a genetic resource somehow imbue and embody the chemical and genetic properties of that physical material to the extent that they are inherent within the definition of that physical resource. DSI is simply an abstraction that is representative and informative about a physical material, the genetic resource, but it is not the resource itself.

Any proposal to extend the concept of a ‘genetic resource’ to encompass DSI on that resource is an attempt to extend the terms of the CBD and NP beyond the originally intended and agreed scope of physical materials as articulated in Article 2 of the NP *inter alia*. To extend the CBD and NP into the realm of information control represents a clear departure from the original intentions and aims of the CBD. This presents multiple challenges to wider aspects of society and international relations that is beyond the remit of the *ad hoc* technical expert group (AHTEG) to analyse fully within the time span available.

If a genuine need is shown to exist for DSI to be controlled by a convention similar to the CBD so that it may achieve the three primary objectives, we believe that this should be the subject of an entirely new protocol. It is our view that it is inappropriate to contort the currently accepted definition of ‘genetic

resources' within the CBD and NP to such an extent that the integrity of the CBD and NP as a whole is compromised.

DSI does not constitute “Traditional knowledge associated with genetic resources”

“Traditional knowledge associated with genetic resources” is defined in the NP as traditional knowledge held by an indigenous or local community that is relevant for the utilisation of genetic resources and that is described in the mutually agreed terms applying to the utilisation of genetic resources. As such, DSI cannot be considered as constituting a form of traditional knowledge. If DSI were to be considered to be traditional knowledge presumably DSI generated in a laboratory would belong to the investigators in the laboratory and so, presumably, in many cases to their employer.

Imposition of benefit-sharing arrangements from commercial and non-commercial use of DSI on genetic resources would have a profound effect on freedom of thought and freedom of research

Article 5 of the NP specifies that an intention of the CBD is to ensure that benefits arising from the utilization of genetic resources as well as subsequent applications and commercialization shall be shared in a fair and equitable way with the Party providing such resources. Free sharing of information, such as DSI, represents a fundamental benefit in its own right. Indeed, the free dissemination of DSI arising from physical genetic resources that have been subject ABS on MAT represents in its own right the very manifestation of the objectives of Article 5.

Article 17 of the CBD requires that Parties shall facilitate the exchange of information, from all publicly available sources, relevant to the conservation and sustainable use of biological diversity, taking into account the special needs of developing countries. The CBD, therefore, explicitly promotes the exchange of information, of which DSI is one type.

Parties, other Governments, indigenous peoples and local communities are the fundamental beneficiaries of a world where free and unfettered use of DSI allows for the production of new medicines, improved crops, new materials, biohazard monitoring and bio-vigilance to name but a few.

It is practically impossible to divide commercial from non-commercial use of DSI. The gap between the academic and private sectors is not clear cut. A great many of the world's major academic institutions from a great many countries are some of the largest holders of patented technologies. These intellectual property estates are a vital source of income to support fundamental basic research that benefits all of humanity. Hindering the public-funded sector from the ability to leverage their research base for commercial objectives, through the imposition of additional regulatory burdens such as ABS for DSI, could lead to catastrophic loss of interest in, and of vital funding streams for, that research.

Imposing a regime of benefit-sharing on MAT for DSI could result in either widespread non-compliance or, more worryingly, a mass movement of important research away from DSI on genetic resources that are subject to NP regulatory burdens. CIPA has been informed from other major stakeholders in the life science industry that such movement away from research on physical genetic resources subject to ABS is already occurring. In at least one example that we are aware of, following due diligence one major biotech company, which had previously been researching on “out of NP scope” material, opted not to pursue a similar research programme on material that was “in NP scope” due to a lack of clarity of the administrative process in the country of origin. If the definition of genetic resources is extended to include DSI it is likely to introduce multiple further parties from whom MAT and PIC will have to be obtained so increasing the difficulties. We fear that this movement could accelerate significantly. Preventing or reducing the sustainable utilisation of important genetic resources and the transfer of technology is contrary to the objectives of the CBD. It cannot be the intention of the CBD to frustrate beneficial research and transfers.

Widespread non-compliance with expanded scope of the NP to include DSI, could lead towards criminalization of legitimate commercial and non-commercial research in countries that are Party to the NP. Once again, this could lead to outcomes that are fundamentally in contravention with the objectives of the CBD, most notably the conservation of biodiversity, and the sustainable use of its components. Hence, imposition of MAT for DSI could frustrate the three primary objectives of the Convention: the conservation

of biodiversity, the sustainable use of its components and the fair and equitable sharing of the benefits arising out of the utilization of genetic resources.

The CBD recognises that some measures, policies or practices induce behaviour that is harmful for biodiversity, often as unanticipated side effects to policies designed to attain other objectives. Such unintended consequences are referred to in the CBD as ‘perverse’ incentives. Whilst the CBD clearly views subsidies that promote overproduction or consumption of resources as perverse incentives, it is important to recognise that denial of free access to DSI and implementation of tariffs or controls on the use of information could also have a perverse effect on biodiversity and sustainable use of genetic resources.

Imposition of benefit-sharing arrangements on use of DSI could restrict access to information for the purpose of conservation of biological diversity and the sustainable use of its components

An unintended consequence of the proposals to include DSI within the remit of the NP could be to limit access to DSI for researchers around the world. The vast majority of biotechnology occurs within commercial and non-commercial R&D facilities within those Parties, or other Governments (such as the United States of America), that have developed advanced economies. For decades these advanced economies have utilized genetic resources, in the process of which they create and will continue to create DSI that could fall within a proposed remit of the NP.

Imposing a regime of benefit-sharing on MAT for DSI could result in advanced economies electing to control free access to DSI originating within their biotechnology industries and research facilities. In this way, the NP might serve to empower advanced economies to the detriment of conservation of biological diversity and the sustainable use of its components. In particular, loss of access to DSI might serve to impoverish developing nations, indigenous peoples and local communities by depriving them of the free access to the use of DSI they currently enjoy. Hence, an unintended consequence of extending the NP to cover use of DSI might be to hand the advanced economies yet a further mechanism to exert advantage over developing nations.

At present the benefits arising from the utilization of genetic resources as well as subsequent applications and commercialization that is vested in the DSI created by these advanced economies is generally shared by all Parties that can access them online. In effect, DSI is by-and-large considered an open resource. This benefit is available to communities in all Parties to the NP and beyond, including indigenous peoples and local communities.

Article 16 of the CBD recognises that access to, and transfer of, technology among Parties are essential elements for the attainment of the objectives of the CBD. It is apparent that creating a regime that allows advanced economies to control access to DSI until MAT have been agreed will be disadvantageous to developing nations who have limited resources available.

It is not clear from any proposals so far put forward whether ABS resulting from the use of DSI would be allocated directly to the objectives of the CBD or would remain within the remit of the originator country. Allowing nations to control the use and exploitation of DSI, under the umbrella of the CBD, risks affording a sovereign moral right to the developed nations to levy the rest of the world for the use of DSI. This use might be in the form of the information itself, or more worryingly in the form of products, perhaps even medicines. This cannot possibly be to the benefit of the developing world or its indigenous communities.

It is our position that any measures taken under the CBD that restrict the use and exploitation of DSI will correspondingly reduce the dissemination of knowledge, contrary to the aims of the CBD.

DSI should not be included within the remit of the CBD or Nagoya, and nor should a convention be introduced which attempts to impose similar controls on DSI.

We at CIPA respectfully endorse these views.

EcoHealth Alliance**SUBMISSION**

EcoHealth Alliance maintains its support for the CBD's efforts to ensure shared benefits and intellectual property. On the issue of digital sequence information (DSI), we encourage Delegates to establish internationally-harmonized mechanisms that provide clear and consistent processes and promote timely information sharing to avoid unintended consequences that have been observed in related contexts (e.g., recent delays in sample permitting processes that impeded investigations of wildlife mass mortality events that threatened entire populations or subpopulations).

A standard form similar to a standardized Material Transfer and Intellectual Property Agreements, agreed by parties, would reduce the process variation among countries or parties that can cause inadvertent delays. One component of this agreement should address data sharing arrangements between parties and research entities and consortia operating in the country to ensure built-in processes and expectations for DSI review and release. For example, consistent with ABS as well as a One Health approach, the USAID Emerging Pandemic Threats PREDICT project together with its government partners developed data sharing processes in 30 countries, under which project data (including DSI) are reviewed, interpreted and approved for public release by authorities responsible for environmental, animal, and human health while maintaining pre-established property rights and benefits for all parties.

To this end, we note the efforts of the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) to establish a simplified procedure for the international movement of emergency diagnostic specimens from species of conservation concern. Particularly in the case of disease investigation, DSI is requisite for effective detection and diagnosis of pathogens and pests both known and novel, where international scientific collaboration is critical for informing appropriate health and conservation-based measures. Furthermore, this sharing is critically important to understand the extent of microbial diversity as well as the species vulnerable to infection and disease, including their transboundary movement, so as to establish a baseline and inform appropriate prevention and control. These measures are wholly in line with the promotion of the Aichi Biodiversity Targets.

EcoHealth Alliance wishes to thank the Secretariat for the opportunity to comment on this important topic.

NOTE: Submitted on behalf of EcoHealth Alliance by: Catherine Machalaba, Policy Advisor and Research Scientist, EcoHealth Alliance

EuropaBio - The European Association for Bioindustries**SUBMISSION**

EuropaBio files this submission as a response to the Convention on Biological Diversity (CBD) decision 14/20, as well as in support of the position of the International Chamber of Commerce responding to the same call.

EuropaBio recommends the substitution of the term digital sequence information (DSI) by the more precise term "Genetic Resource Sequence Data" (GRSD), to facilitate a more fact- and science-based discussion on this subject matter. Furthermore, imposing additional ABS obligations – other than through the existing mechanism of mutually agreed terms - on access, use and/or dissemination of GRSD, and any more open ended concept, including DSI, would have a significant negative impact on the future of biotech research and the benefits resulting from it for society, as further outlined below. EuropaBio takes the view that over the years the term DSI has taken different meanings in the relevant CBD and/or the Nagoya Protocol discussions and the term is, therefore, qualified by many as unclear and imprecise. We side with the view of the international business representation that the term "Genetic Resource Sequence Data" or "GRSD" is clearer and more precise from a scientific and legal point of view.

The term GRSD refers to the description of the order of nucleotides (DNA or RNA), as found in nature, in the genome or encoded by the genome of a given genetic resource. The “genome” includes nuclear and extra-nuclear DNA, and coding (gene) and non-coding DNA sequences³⁷. It does not include other molecules resulting from natural metabolic processes associated with or requiring the genome.

In the term GRSD “Genetic Resource” refers to a term already defined in the CBD and provides for the necessary link with the DNA or RNA sequence that provides an exact description of the DNA or RNA in a specific genetic resource, thereby ensuring the direct link with the genetic and/or biochemical composition of the genetic resource. GRSD does not and cannot include a DNA or RNA sequence not identical to that found in a genetic resource. In addition, GRSD must originate from a physical source, consistent with the CBD definition of ‘genetic resource’. “Sequence” results from the process of determining the order of nucleotides in a DNA genome or RNA molecule of a genetic resource of a specific species. It refers to the order of nucleotides in the genome. “Data” refers to the actual genetic sequence data of a specific genetic resource. GRSD cannot, and does not, include information connected with or resulting from the analysis or further application of GRSD, e.g. sequence assembly, sequence annotation, genetic maps, metabolic maps, three-dimensional structure information or physiological properties related to it.

To the extent permitted under the existing CBD and Nagoya Protocol legal framework, benefit-sharing provisions relating to GRSD may already be included by provider countries in mutually agreed terms with regard to the utilization of a specific genetic resource.

Any attempt to amend the scope of the CBD and the Nagoya Protocol to allow for the imposition of additional ABS obligations on access, use and/or dissemination of GRSD³⁸ will negatively affect biotech research and the societal benefits therefrom, such as the availability of innovative products improving food security and human health. A system of open exchange of GRSD for advancing research and development will help achieve the objectives of the CBD and the Nagoya Protocol by enabling the sustainable use and hence supporting the valorisation and thus conservation of biodiversity. Specific reference is made to the fact that there is currently a well-established and functioning international framework supporting the open exchange of GRSD, consistent with established principles of ethical and responsible scientific research; and that considerable “non-monetary” benefits are derived by all countries from the open exchange of GRSD.

EuropaBio remains committed to the objectives of the CBD and the Nagoya Protocol. Our members are active users of genetic resources, make large investments in research and development to unlock the potential of genetic resources and favour the creation of a proportionate regulatory environment to enable their continued characterisation, conservation and sustainable use, thereby safeguarding the many societal benefits this provides

Finally, we are committed to providing additional insights on the well-functioning of the current system of open exchange of data by supporting detailed case studies by the international business community.

About EuropaBio

EuropaBio, the European Association for Bioindustries, promotes an innovative and dynamic European biotechnology industry. EuropaBio and its members are committed to the socially responsible use of biotechnology to improve quality of life, to prevent, diagnose, treat and cure diseases, to improve the quality and quantity of food and feedstuffs and to move towards a bio-based and zero-waste economy. EuropaBio represents 79 corporate and associate members and bio-regions, and 17 national biotechnology associations in turn representing over 1800 biotech SMEs. Read more about our work at www.europabio.org.

³⁷ In the case of RNA viruses the genome consists of RNA instead of DNA.

³⁸ The concerns expressed regarding the negative impact of imposing additional ABS obligations applies a fortiori to every more open-ended term, including “DSI”.

International Barcode of Life Consortium (iBOL)**SUBMISSION**

The International Barcode of Life Consortium (iBOL) is pleased to submit the following comments in regard to the topic of notification 2019-012 “Digital Sequence Information on Genetic Resources: Submission of views and information and call for expression of interest to undertake studies”.

We recognize that the closing date for submissions to this notification was 30 June 2019 but, in light of the launch of iBOL's new seven-year program, BIOSCAN, we believe that it is timely for us to provide these comments in the hope that they may be of value to the ad hoc technical expert group (AHTEG).

We recognize the difficulty faced in resolving an appropriate definition for Digital Sequence Information (DSI) in the context of the Access and Benefit Sharing (ABS) arrangements under the Nagoya Protocol, and also understand the concern from some Parties over delays arising from continued discussion on this point. Nevertheless, we believe that it is important for the AHTEG and for the Conference of the Parties to understand and consider the full range of possible uses of genomics technologies and of data derived from genetic resources in the light of all three objectives of the Convention.

We note the recommendation of several respondents to the notification that the problem may be addressed by considering DSI as a continuum from raw sequence data to patentable inventions. We would add that there is a further related continuum from short fragments of sequence data, particularly for non-coding regions of the genome, through to fully annotated complete genomes. Both of these dimensions, as well as the possible uses of each application of genomic technologies, need to be under consideration in determining the scope of DSI for ABS.

To assist with consideration of these dimensions of variation, we wish to highlight relevant characteristics of the DNA barcoding and metabarcoding solutions developed and promoted by iBOL.

The relevance of DNA barcoding to Conservation and Sustainable Use of Biodiversity has already been recognized in the Note by the Executive Secretary, CBD/DSI/AHTEG/2018/1/3 (especially paragraph 30ff).

More specifically, the proven and developing applications of DNA barcoding in areas of key public concern include at least the following³⁹:

- Invasive and alien species – identifying and monitoring invasive organisms and their ecological impact, improving early detection and regulatory measures to curb cross-border transfer of alien species;
- Endangered species – enhancing taxonomic and ecological knowledge about endangered species and creating a diagnostic framework for monitoring and curbing illegal harvest and trade through improving forensic approaches and streamlining regulatory frameworks;
- Agriculture and forestry – identifying and monitoring agriculture and forestry pests and biological control agents;
- Human health – identifying and monitoring human disease vectors and reservoirs; reconstructing disease transmission pathways; assessment and monitoring of natural-borne disease foci;
- Environmental surveillance/monitoring – helping extractive industries (e.g. oil, gas, mining), the natural resources (forestry, fisheries) and agriculture sectors to meet their environmental compliance requirements and to evaluate the efficiency of offset, restoration and remediation measures;
- Market surveillance, product ingredient authentication; detection of food contamination and substitution (e.g. seafood, meat and natural products).

³⁹ Davis, K & Borisenko A, 2017, Introduction to Access and Benefit-Sharing and the Nagoya Protocol: What DNA Barcoding Researchers Need to Know. <https://doi.org/10.3897/ab.e22579>

iBOL's first program, BARCODE 500K (2010-2015)⁴⁰ engaged stakeholders from all regions in constructing a reference library of short DNA sequences for use in identification of more than 500,000 species. BIOSCAN (2019-2026)⁴¹ aims to build on this foundation by expanding the reference library to include at least 2 million species and by operationalizing the use of DNA barcodes for low-cost rapid surveys and continuous monitoring of biodiversity. BIOSCAN aims to use this approach to gain understanding of biodiversity composition at thousands of sites across different ecosystems around the planet.

DNA barcoding uses short genetic markers that represent either gene fragments, e.g. ribulose biphosphate carboxylase large chain (rbcLa) in plants and cytochrome c oxidase subunit I (COI) in animals, or nonfunctional elements of the genome, e.g., internal transcribed spacer (ITS1) in fungi. These markers are selected because they show rapid changes as species and populations diversify. This makes them ideally suited for species-level identifications, but these sequences have limited value except as part of a system of DNA barcodes to support species identification.

DNA barcoding has relevance to many aspects of the mission of the CBD but most significantly it provides an efficient and cost-effective way to address the taxonomic impediment. The greatest challenge to developing a comprehensive baseline understanding of the diversity of life on Earth and to understanding the geographic range, status and trends for every species is the difficulty, cost, and time required to identify most groups of organisms. DNA barcoding offers taxonomists fundamental data on the diversity found in any group at any location and supports their efforts to describe currently unnamed species. Even where expertise is lacking to complete the work of species description, it allows unnamed units of species diversity to be mapped and monitored across time and space. As the reference library of DNA barcodes becomes more complete, metabarcoding approaches facilitate low-cost detection of whole communities of species from environmental samples.

It is realistic to expect that observation and monitoring of biodiversity patterns for conservation, food security, and global sustainability will soon be possible primarily through a combination of DNA-based techniques, particularly barcoding and metabarcoding, and of remote-sensing. DNA barcoding will accordingly be a significant component in the earth-observing system for biodiversity that is necessary to support the Post-2020 Global Biodiversity Framework and the Sustainable Development Goals.

Achieving these benefits will depend on collaboration at a global scale to complete the DNA barcode reference library and to develop monitoring protocols applicable to all ecosystems. iBOL, through BIOSCAN, is working to deliver these components. However, this whole approach depends for its effectiveness on the comprehensiveness and openness of the data included in the reference library. Open sharing of these sequences by all Parties will bring significant benefits to everyone. Global participation will reveal the species or varieties that are unique to each country or region and will assist with development of National Biodiversity Strategies and Action Plans (NBSAPs) and with assessing the coverage and completeness of the world's system of protected areas.

iBOL is working to accelerate the application of genetic markers as a global solution for the good of all nations and for improved management of biodiversity. This includes a commitment to support national communities of practice as members of a global consortium and to develop materials and solutions to enhance capacity and technical infrastructure in all regions.

We urge the AHTEG and all Parties to consider and safeguard public-good applications of DSI such as DNA barcoding and metabarcoding within the framework for Access and Benefit Sharing under the Nagoya Protocol.

We welcome any opportunities to provide further information or clarifications to assist the work of the AHTEG.

⁴⁰ <https://ibol.org/programs/barcode-500k/>

⁴¹ <https://ibol.org/programs/bioscan/>

International Chamber of Commerce (ICC)

SUBMISSION

Key messages:

1. ICC proposes using the term “Genetic Resource Sequence Data” (GRSD) instead of “Digital Sequence Information” (DSI) to enable a more fact- and science-based discussion.
2. Open access to and use of DSI for advancing research and development creates huge benefits, both from commercial and non-commercial use, which help achieve the objectives of the United Nations Convention on Biological Diversity (CBD) and the Nagoya Protocol.

This submission is provided in response to Convention on Biological Diversity (CBD) decision 14/20, paragraph 9, which invites parties and stakeholders to provide views and information on:

- relevant terminology and scope related to “Digital Sequence Information” (DSI) and
- benefit sharing from commercial and non-commercial use of DSI.

Relevant terminology and scope related to “Digital Sequence Information”

When the term “DSI” arose in CBD discussions, it was understood to refer to electronically stored and exchanged DNA sequence information. The content of the term has since expanded, and discussions on this subject matter are now confounded by multiple interpretations of it.

The term “DSI” has been qualified by many as being unclear and imprecise. The report of the Ad Hoc Technical Expert Group on DSI states that there is agreement that the term “DSI” is not apt and should be used as a mere placeholder. ICC agrees with the view that the term “DSI” is imprecise and not appropriate for the ongoing discussions.

ICC would like to propose instead the term “**Genetic Resource Sequence Data**” or “GRSD”, as a clearer, more scientifically-precise alternative which would be more appropriate for use in the discussions. ICC is proposing this term solely to facilitate a more fact- and science-based discussion on this subject matter.

Genetic Resource Sequence Data refers to the description of the order of nucleotides (DNA or RNA), as found in nature, in the genome or encoded by the genome of a given genetic resource. The “genome” includes nuclear and extra-nuclear DNA, and coding (gene) and non-coding DNA sequences. It does not include other molecules resulting from natural metabolic processes associated with or requiring the genetic resource.

1. “Genetic Resource” provides for the necessary link with the DNA or RNA sequence that provides an exact description of the DNA or RNA in a specific genetic resource, thereby ensuring the direct link with the genetic and/or biochemical composition of the genetic resource. It makes a clear reference to the fact that the genetic sequence correlates exactly with the DNA or RNA in a specific genetic resource. In the context of the discussions under the CBD and/or the Nagoya Protocol, GRSD does not and cannot include a DNA or RNA sequence that is not identical to that found in a genetic resource. In addition, it is important to note that GRSD must originate from a physical source, consistent with the definition of ‘genetic resource’ in the CBD.
2. “Sequence” results from the process of determining the order of nucleotides (nucleic acids) in a DNA genome or RNA molecule of a genetic resource of a specific species.
3. “Data” refers to the actual genetic sequence data of a specific genetic resource. Data is to be distinguished from information since “information” involves a level of additional processing or analysis. Thus, GRSD cannot, and does not, include information connected with or resulting from the analysis or further application of GRSD, e.g. sequence assembly, sequence annotation, genetic maps, metabolic maps, three-dimensional structure information or physiological properties related to it. Including information resulting from human interaction on GRSD would result in yielding man-made genetic sequences, which would no longer be considered GRSD.

Regarding “Digital Sequence Information”, our understanding is that the word “digital” refers to the format in which nucleic acid sequences may be stored and shared (e.g. in a sequence database) and used for subsequent analyses or uses. ICC does not consider this necessary to retain as this does not describe the genetic sequence itself.

ICC is proposing the term “Genetic Resource Sequence Data” to enable a more fact- and science-based discussion. This does not impact ICC’s position, as communicated in its previous submissions on DSI and in the joint statement by numerous stakeholder organisations on the need to maintain open exchange of DSI.⁴²

“Digital Sequence Information” & existing systems of benefit sharing

As pointed out in the section above, ICC has proposed the term “Genetic Resource Sequence Data” or “GRSD” for use in the ongoing discussions on this subject matter. “GRSD” is hence used in this section on the existing systems of benefit sharing.

Under the existing CBD and Nagoya Protocol framework, benefit-sharing provisions relating to GRSD can already be included by provider countries in mutually agreed terms related to the utilisation of a specific genetic resource. However, the subject of obligations under a MAT should not include open-ended or ambiguous concepts such as “DSI”.

ICC wishes to emphasise that amending the scope of the CBD and the Nagoya Protocol to allow for the imposition of additional Access and Benefit Sharing (ABS) obligations on access to and/or the use of GRSD⁴³ - other than through the existing mechanism of mutually agreed terms - will have a significant negative impact on the future of biological research and the benefits resulting from it for society. Open access to and use of GRSD for advancing research and development creates huge benefits, from both commercial and non-commercial use, which help achieve the objectives of the CBD and the Nagoya Protocol. ICC would like to explicitly emphasise the following key elements to be carefully considered:

1. GRSD is a critical tool in the conservation and sustainable use of genetic resources, and its use is currently leading to many societal benefits, including the availability of innovative products improving food security and human health. Impeding the flow and use of data or information would work against research projects in several different areas, many of which contribute to the objectives of the CBD and the Nagoya Protocol. The CBD specifically encourages the exchange of information to support conservation, protection and sustainable use of biodiversity.⁴⁴ The system of open exchange of GRSD is essential to unlock the value of physical genetic resources, by enabling the sustainable use and hence supporting the valorisation and thus conservation of biodiversity. One of the original aims of the Nagoya Protocol was to incentivise countries and biodiversity custodians to preserve physical ecosystems and species by providing a mechanism which would allow them benefit from such conservation. This is clearly stated in the preamble of the Nagoya Protocol by “*Recognizing that public awareness of the economic value of ecosystems and biodiversity and the fair and equitable sharing of this economic value with the custodians of biodiversity are key incentives for the conservation of biological diversity and the sustainable use of its components*”. If the focus of benefit sharing shifts to data, any incentives provided by the ABS mechanism to preserve physical ecosystems and species will be greatly reduced. It is therefore of paramount importance that GRSD which is now in the public domain continues to remain freely accessible to achieve the broader aims of the CBD.
2. Making the sharing of GRSD more legally uncertain, time-consuming, administratively burdensome and costly by regulating its use conflicts with the aim of the Nagoya Protocol to

⁴² Promoting sustainable use and conservation of biodiversity through open exchange of Digital Sequence Information, 8 November 2018

⁴³ The concerns expressed regarding the negative impact of imposing additional ABS obligations applies *a fortiori* to every more open-ended term, including “DSI”

⁴⁴ Article 17, Convention on Biological Diversity

improve legal certainty and will have a chilling effect on the use of GRSD in biological research, and consequently lead to less benefits. All countries rely on and benefit from unencumbered access to and exchange of GRSD to deal with key societal issues, such as human, animal and plant health, food security and the environment. There is currently a well-established and functioning international framework supporting the open exchange of GRSD, consistent with established principles of ethical and responsible scientific research that are foundational for the advancement of science. Open exchange is especially important for GRSD, the unencumbered use of which allows the swift compilation, comparison and re-analysis of information from a variety of sources, across multiple databases and gene sequences.

3. Considerable “non-monetary” benefits can be derived by all countries from the open exchange of publicly available GRSD. The extensive public GRSD database managed by the International Nucleotide Sequence Database Collaboration (INSDC), maintained at the expense of the three host governments, is freely accessible to researchers in all countries; and was accessed by 172 countries from all regions between 2014 and 2016.⁴⁵ Open exchange of GRSD is also necessary for international research collaborations, which not only allow the pooling of expertise and resources to resolve problems of global or regional relevance but are also essential vehicles for capacity building and the exchange of knowledge and expertise. Countries also benefit from the technologies and products resulting from research supported by open exchange of GRSD, wherever this research takes place. Widespread regulation of GRSD is likely to result in a dramatic reduction in information being made available in public databases. There is therefore a danger that the significant non-monetary benefits which are currently being delivered to developing countries could be diminished in the process. Considering the number of accessions from public databases by developing countries, small research institutions in developing countries would be especially burdened by additional obligations.
4. Numerous legal interpretations have confirmed that the definition of genetic resources refers to tangible material and does not include immaterial information. Negotiations to change this definition would require years, or even decades, as well as resources that would be better spent on efficient implementation of effective and workable measures that achieve the three objectives of the CBD, including capacity building to ensure sustainable use.

In short, imposing ABS obligations on the access, use and/or dissemination of GRSD, other than through the existing mechanism of mutually agreed terms, would substantially jeopardise the creation and sharing of benefits, which contribute to achieving the objectives of the CBD and the Nagoya Protocol, and ultimately support managing some of society’s biggest challenges such as food security and human health.

In order to provide additional insights on the well-functioning of the current system of open exchange of data, ICC is currently also working on detailed case studies, which will be provided in due course to further enhance the informed discussions on this topic in the context of the CBD and the Nagoya Protocol.

⁴⁵ “As a member of INSDC (International Nucleotide Sequence Database Collaboration) maintaining the international GRSD database, we found that a total of 172 countries were using the database during the period from 2014 to 2016 (Europe 46, Asia 43, South/Central America 35, Africa 32, Oceania 14, and North America 2; Website access: 1,621,300) upon investigating the state of the use from outside Japan of DDBJ (DNA Data Bank of Japan) operated by National Institute of Genetics”- *Extract from the Submission by the Government of Japan on the “Current state of the use of digital sequence information on genetic resources in the biodiversity field”*.

International Fragrance Association (IFRA) and International Organization for the Flavor Industry (IOFI)**SUBMISSION**

IFRA and IOFI are pleased to contribute to the request for information launched by the Secretariat of the Convention on Biological Diversity (CBD) with the decision 14/20, paragraph 9, which invites parties and stakeholders to provide views and information on:

- relevant terminology and scope related to "Digital Sequence Information" (DSI) and
- benefit sharing from commercial and non-commercial use of DSI.

What is currently described as DSI is an important source of information for many Flavour and Fragrance companies with biotechnology related R&D activities. Biodiversity is a highly valuable source of innovation and inspiration for new flavourings, fragrances, or other ingredients. For this reason, IFRA and IOFI are strongly committed to the principles of the UN Convention on Biological Diversity and the Nagoya Protocol on Access to Genetic Resources and the Equitable Sharing of Benefits arising from their utilization.

So-called DSI is not covered by the obligations of the Nagoya Protocol as it stands. Amending the scope of the CBD and/or the Nagoya Protocol to allow for the establishment of additional Access and Benefit Sharing (ABS) obligations on access to and/or the use of DSI (other than to rely on the already existing mechanism of PIC and MAT, which are currently only mandatory for physical genetic resources) will have a significant negative impact on the future of biological research and the benefits resulting from it for society.

It is not clear how, for DSI, compliance with the principles of the Nagoya Protocol and related regulatory obligations can be ensured, given that data and information on origin, providers, or sovereign right holders of the initial GR are usually not included in or connected with respective DSI databases. Indeed, many species share a large proportion of their genetic sequence so it is difficult to attribute a sequence to a particular species.

Moreover, the inclusion of DSI under the scope of the Nagoya Protocol will place an enormous administrative burden on bio-based research. Open access to and use of DSI for advancing research

¹ The International Fragrance Association (IFRA) and the International Organization for the Flavor Industry (IOFI) respectively represent the fragrance and the flavor industry worldwide. IFRA and IOFI membership comprises direct company members as well as national associations in Europe, North America, South America and Asia Pacific. More information can be found at www.ifrafragrance.org and at www.iofi.org.

and development creates huge benefits, from both commercial and non-commercial use, which help achieve the objectives of the CBD and the Nagoya Protocol. The lack of legal certainty that would be created with the inclusion of DSI under the scope of the Nagoya Protocol will likely discourage potential users to invest in bio-based innovations, thus reducing the potential for benefit creation and sharing.

IFRA and IOFI would kindly ask the CBD Secretariat to consider the above-mentioned observations / concerns. IFRA and IOFI are fully supportive of the position of the International Chamber of Commerce (ICC) (See attached), including their recommendations on:

- The use of the term "Genetic Resource Sequence Data" (GRSD) instead of DSI to enable a more fact- and science-based discussion.
- The fact that open access to and use of DSI for advancing research and development creates huge benefits, both from commercial and non-commercial use, which help achieve the objectives of the United Nations Convention on Biological Diversity (CBD) and the Nagoya Protocol.

IFRA and IOFI remain at your disposal in case of further information request.

International Nucleotide Sequence Database Collaboration (INSDC)

SUBMISSION

Key points

- INSDC databases provide the long-established and broadly adopted data infrastructure for the open sharing of sequences
- Open access to data is a central element of INSDC operations and lies under the governance of governments, institutions and scientific peers
- Submission to INSDC databases is an essential scientific process that adds value and drives knowledge generation from the entire corpus of data
- Use of INSDC databases is embedded in the scientific process, both for data submission and data access
- INSDC provides the effective free and open data infrastructure required to enable biodiversity-related Access and Benefit Sharing

Terminology and Scope

The International Nucleotide Sequence Database Collaboration (INSDC; <http://www.insdc.org/>) has for four decades captured, curated and preserved the world's output of DNA and RNA sequence data, including the sequences themselves (ordered strings of letters that represent the order of the nucleotides on chromosomes) and annotations thereon (such as indications of genes and their functions).

These data are variously referred to as "Digital Sequence Information; DSI" and "Genetic Sequence Data; GSD". These are data and do not constitute genetic material; rather they are data that describe genetic material.

Access to GSD and benefit-sharing

INSDC is a partnership between three major institutions, the National Institute of Genetics' DNA Data Bank of Japan (DDBJ), the European Molecular Biology Laboratory's European Bioinformatics Institute (EMBL-EBI) and the US National Institutes of Health's National Library of Medicine, National Center for Biotechnology Information (NCBI), who provide globally comprehensive coverage of open access data assured through data exchange systems and their supporting technical data standards. The open access policy (see below for details) is dictated by governments, funding sources, institutional governance and the INSDC International Advisory Committee, for the use of scientists worldwide. INSDC databases do not set these policies and could not make changes to suit any restrictions on open access that might be imposed by the Convention or the Nagoya Protocol.

The process of data submission into an INSDC database is one in which the data provider structures and describes the new data set in a standard manner, such that it can be integrated into the entire corpus of INSDC content. Data providers are encouraged to provide contextual metadata which describes when and where the biological sample was obtained and where it is stored, if applicable (for example, a museum voucher). Importantly, it is through this integrated corpus of data that the fullest value can be extracted; few scientific interpretations rely on a single sequence and most rely on the aggregate of many different sequences considered together.

Submission to INSDC databases is broadly accepted by the scientific community as an essential part of the process of carrying out scientific work. Mandatory submission to INSDC is required prior to, or at the time of, publication in the academic literature in the vast majority of life science journals covering the life sciences. Equally, use of the data within INSDC databases is an accepted, and often daily, norm for life scientists. INSDC has thus long been adopted by the community as the system for managing sequence data.

INSDC is supported mainly by the host governments to promote and support open science primarily for human genetic research, food security and public health. Biological research, including biodiversity research, also benefits from this investment. While sequences that are CBD-relevant (in terms of legal

scope) are substantial, but a minor part of the database (relative to human genome sequences or model organisms), these sequences play an essential enabling role in meeting the first two objectives of the CBD, the conservation of biological diversity and the sustainable use of its components.

Open Access in INSDC

Services provided by INSDC databases are free of charge to all users, both those submitting data and those searching and retrieving data.

Data presented in INSDC are open access and available freely across a host of online services, such as web sites, search tools, programmatic interfaces and FTP sites. It is only with open access that data can be fully integrated to support interpretation by original providers and the broader scientific community.

The INSDC partner databases place no conditions on the use and redistribution of data and records remain within the ownership of the data submitters; INSDC database are the hosts of the data. INSDC databases Terms of Use state that the databases provide data openly, but note that data owners or third parties may assert conditions on the use or application of sequences relating to rights such as intellectual property and access and benefit sharing (see <https://www.ddbj.nig.ac.jp/policies.html#credit>, <https://www.ebi.ac.uk/about/terms-of-use>, <https://www.ncbi.nlm.nih.gov/home/about/policies/#data>)

The INSDC partners believe that open access data are a key component of any successful Access and Benefit Sharing system. Without open access and data integration, the value (and ultimate benefit) from the data will be significantly reduced. Regulation of GSD would increase barriers to innovation and discovery which would affect attaining the goals of CBD as well as impacting food safety and health worldwide.

Illustrative Examples - Benefits of Open Genetic Sequence Data

Genomic sequence analysis is being used in the public health arena. In late 2017, there was an outbreak of Salmonella in the US. The genomes of bacteria isolated from clinical samples were sequenced and analyzed. Genomes were clustered to determine which are most closely related and potentially derived from a single source. The sequences from this outbreak clustered with a Salmonella genome isolated from a papaya which originated in Mexico (<https://www.cdc.gov/salmonella/kiambu-07-17/>). Open sequence data allowed public health scientists to determine the cause of the outbreak and prevent additional illness.

Open genetic sequence data enabled the discovery of fish sold in markets and restaurants may be mislabeled. DNA sequence can distinguish different species that may look very similar to the untrained eye. The DNA sequence of a gene from restaurant and fish market samples were compared with fish gene sequences in INSDC to determine the species of the sample. The species indicated was not always the species of fish that was detected by DNA sequencing. For example: (1) critically endangered shark species were being sold as different types of fish (<https://www.nature.com/articles/s41598-018-38270-3>) (2) less expensive fish were sold as more expensive ones (<https://onlinelibrary.wiley.com/doi/full/10.1111/cobi.12888>). This fraud would not have been uncovered without open GSD.

Across INSDC, we enable access and share the added value of integrated data to all. We believe that keeping GSD access open would enable more access and benefit sharing, not less.

NOTE: Signatories: Ilene Karsch-Mizrachi (NCBI), Guy Cochrane (EMBL-EBI), Yasukazu Nakamura (DDBJ) on behalf of the INSDC.

Japan Bioindustry Association (JBA)
--

SUBMISSION

Concerning the notification 2019-012 (SCBD/NPU/DC/VN/KG/RKi/87804) of 5 February 2019, Japan Bioindustry Association (JBA) greatly appreciates the opportunity for submission of our views and relevant information on the issues of digital sequence information on genetic resources, as a relevant organization.

We would be most grateful if the Executive Secretary would take into consideration our views and relevant information given below, in the process of preparing documentation on the related agenda item for the meeting of the extended Ad Hoc Technical Expert Group.

1. Background

The fourteenth meeting of the Conference of the Parties to the Convention on Biological Diversity (hereafter the CBD) and the third meeting of the Conference of the Parties serving as the meeting of the Parties to the Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefit Arising from their Utilization to the Convention on Biological Diversity (hereafter the Nagoya Protocol), each adopted the decisions on digital sequence information on genetic resources (decisions 14/20 and NP-3/12, respectively). In particular, decision 14/20 invites Parties, other Governments, indigenous peoples and local communities, and relevant organizations and stakeholders to submit views and relevant information to the Executive Secretary on several issues relating to digital sequence information on genetic resources (paragraph 9(a) and (b)).

2. JBA's General Views on Digital Sequence Information on Genetic Resources under the CBD and the Nagoya Protocol

JBA supports the three objectives of the CBD and the objective of the Nagoya Protocol and has been actively engaged in contributing to the proper implementation of these objectives. And we believe that the legal certainty is essential for the implementation of those objectives of the CBD and the Nagoya Protocol.

In our September 8, 2017 submission to you, we presented our views (which have not changed since then) on the relationship between digital sequence information on genetic resources and the CBD/ the Nagoya Protocol. They are reproduced here again for the present discussion:

“First of all, the definition of “genetic resources” is “genetic material of actual or potential value” with “genetic material” being defined as “any material of plant, animal, microbial or other origin

containing functional units of heredity” (see Article 2 of the CBD). Furthermore, these definitions are also applied to the Nagoya Protocol (see Article 2 of the Protocol). Intangible digital sequence information on genetic resources does not fall within the definition of “genetic resources” under the CBD and the Nagoya Protocol, because “genetic resources” are tangible materials.

However, benefit-sharing relating to digital sequence information on genetic resources obtained by utilizing accessed genetic resources is not excluded from the CBD and the Nagoya Protocol. The conditions, including those for benefit-sharing, for dealing with such digital sequence information could be covered by the mutually agreed terms (MAT) between the provider and the user of the genetic resources.”

Please note that Decision 14/20, paragraph 6, mentions that “ (COP) Further notes that, as there is a divergence of views among Parties regarding benefit-sharing from the use of digital sequence information on genetic resources, Parties commit to working towards resolving this divergence through the process established in the present decision, with the aim of strengthening the fulfillment of the third objective of the Convention and Article 15, paragraph 7, without prejudice to the circumstances to which this article applies;”.

Also note that Article 15, paragraph 7 of the Convention stipulates that “...with the aim of sharing in a fair and equitable way the results of research and development and the benefits arising from the commercial and other utilization of genetic resources with the Contracting Party providing such resources. Such sharing shall be upon mutually agreed terms.”

3. JBA’s Specific Views and Relevant Information

3-1. Terminology and Scope of Digital Sequence Information on Genetic Resources (Decision 14/20, paragraph 9(a))

In our view, as mentioned above, intangible digital sequence information on genetic resources does not fall within the definition of “genetic resources” under the CBD and the Nagoya Protocol. However, benefit-sharing regarding digital sequence information on genetic resources generated from accessed genetic resources in the utilization process could be covered by mutually agreed terms (MAT) between the provider and the user of the genetic resources, based on the framework of access and benefit-sharing (ABS) stipulated by the CBD and the Nagoya Protocol

In order to contribute to resolving the issue of a divergence of views among Parties as described in paragraph 6 of decision 14/20, without prejudice to the circumstances to which Article 15, paragraph 7 of the CBD applies, we would like to provide our proposed terminology and scope for the placeholder terminology “digital sequence information on genetic resources”:

JBA’s proposed terminology: nucleotide sequence data on genetic resources

JBA’s proposed scope: nucleotide sequence data generated by sequence determination of accessed genetic resources in the process of their utilization within the framework of the CBD and the Nagoya Protocol.

3-2. “To Clarify If and How Domestic Measures Consider Digital Sequence Information on Genetic Resources” (Decision 14/20, paragraph 9(a))

In 2017, Japanese government put in place a domestic administrative measure to implement the Nagoya Protocol¹.

Japanese government chose not to require users to obtain prior informed consent (PIC) for accessing genetic resources existing in Japan. The Japan’s measure is for users in Japan to comply with, primarily, Articles 15, 16 and 17 of the Nagoya Protocol.

Concerning “if and how domestic measures consider digital sequence information on genetic resources”(decision 14/20, paragraph 9(a)), the measure stipulates that “information concerning genetic resources such as nucleic acid base sequences” is outside the scope of application, because it does not fall under the category of genetic resources to which the Protocol applies.

3-3. Existing Systems for Use of Nucleotide Sequence Data on Genetic Resources (Decision 14/20, paragraph 9(b))

We have already proposed the terminology “nucleotide sequence data on genetic resources” for the ongoing discussion on this issue (see Section 3-1 of this submission). Therefore, “nucleotide sequence data on genetic resources” is used throughout this section in place of “digital sequence information on genetic resources”.

(1) Open access system for publicly available databases

Open access system for publicly available databases on nucleotide sequences (e.g., International Nucleotide Sequence Database Collaboration (INSDC)) is a form of benefit sharing. In this system, anyone in any country can freely access such databases, and use the information in her/his country for any purpose to derive benefits from it.

Government of Japan found that a total of 172 countries were using the database during the period from 2014 to 2016 upon investigating the state of the use from outside Japan of DNA Data Bank of Japan². (DDBJ, a member of INSDC)

Since scientific and technological advances in key areas of human activities, e.g., health, food-agriculture, and environment, are largely supported by open access to these publicly available databases, regulatory interference with these practices would have seriously negative impact on the

¹ Access and Benefit-sharing Clearing-House of CBD : <https://absch.cbd.int/database/record/ABSCH-MSR-JP-238074>)

² The submission by the government of Japan on the “current state of the use of digital sequence information on genetic resources in the biodiversity field: <https://www.cbd.int/abs/DSI-views/JAPAN-DSI.pdf>.

entire world.

(2) Benefit-sharing system for commercial use of nucleotide sequence data generated from accessed genetic resources in the process of their utilization

In the business sector, “mutually agreed terms” (i.e., contract) on business transaction are generally treated as “trade secret” and therefore kept confidential among only those who are involved in the transaction (e.g., providers and users of genetic resources). To JBA’s knowledge, information on such benefit-sharing contracts for commercial use has not been made available.

(3) Benefit-sharing systems for non-commercial use of nucleotide sequence data on genetic resources

In the public sector, information on benefit-sharing systems, e.g. international scientific collaborations funded by governmental organizations, is publicly available.

The Program of “the Science and Technology Research Partnership for Sustainable Development” (SATREPS) is co-funded by Japan Science and Technology Agency (JST), Japan Agency for Medical Research and Development (AMED) and Japan International Cooperation Agency (JICA). The SATREPS Program promotes international joint research based on the needs of partner countries (mostly developing countries), addressing global issues with the aim to produce research outcomes of practical benefits for local, regional and global societies³.

Such international scientific collaborations have often been used as vehicles for capacity-building/development of developing countries for mutual benefits.

The examples given in Table 1 are extracted from the reports on SATREPS program relating to the use of nucleotide sequence data on genetic resources.

Table 1: Examples of international joint research projects relating to nucleotide sequence data which appeared in the SATREPS Program

Starting Year	Project Title	Partner Country
2016	Breakthrough in Nutrient Use Efficiency for Rice by Genetic Improvement and Fertility Sensing Techniques in Africa (Area: agriculture)	Madagascar
2015	Project for Development of Sericulture Research by Applying Biological Resources and Molecular Genetics (Area: agriculture)	Kenya
2014	Establishment of Cryo-bank System for Vietnamese Native Pig Resources and Sustainable Production System to Conserve Bio-diversity (Area: agriculture)	Viet Nam
2013	Development and Adoption of Latin American Low-input Rice Production System Through Genetic Improvement and Advanced Field-management Technologies (Area: agriculture)	Colombia
	Epidemiological Studies on Animal Protozoan Diseases in Mongolia and Development of Effective Diagnostics Measures (Area: animal health)	Mongolia
	Development of Innovative Research Technique in Genetic Epidemiology of Malaria and Other Parasitic Diseases in Lao PDR for Containment of Their Expanding Endemicity (Area: human health)	Lao People's Democratic Republic

³ The SATREPS Program: <https://www.jst.go.jp/global/english/index.html>

2012	Innovative Bio-production in Indonesia (iBioI): Integrated Bio-refinery Strategy to Promote Biomass Utilization using Super-microbes for Fuels and Chemicals Production (Area: biomass)	Indonesia
	The Project on Rice Research for Tailor-made Breeding and Cultivation Technology Development in Kenya (Area: agriculture)	Kenya
	Diversity Assessment and Development of Sustainable Use of Mexican Genetic Resources (Area: biodiversity)	Mexico
2011	Development of Aquaculture Technology for Food Security and Food Safety in the Next Generation (Area: aquaculture)	Thailand
	Development of Rapid Diagnostics and the Establishment of an Alert System for Outbreaks of Yellow Fever and Rift Valley Fever in Kenya (Area: human health)	Kenya
2010	Development of Internationally Standardized Microbial Resource Center to Promote Life Science Research and Biotechnology (Area: science)	Indonesia
	Development of Crop Genotypes for the Midlands and Mountain Areas of North Vietnam (Area: agriculture)	Viet Nam
2009	Development of Genetic Engineering Technology of Crops with Stress Tolerance against Degradation of Global Environment (Area: agriculture)	Brazil

(Note: The projects presented here were those started in 2016 or earlier. SATREPS Program is currently continued.)

In these projects, the elements of capacity building/development in the use of nucleotide sequence data on genetic resources include the following:

- 1) Basic procedures for access to and generation of nucleotide sequence data on genetic resources and genetic analysis,
- 2) Analysis techniques: expression analysis, functional analysis, phylogenetic analysis, diversity analysis and genotyping, and,
- 3) Practical use of nucleotide sequence data on genetic resources: DNA marker assisted breeding.

JBA fully recognizes the needs for and value of capacity building/development for developing countries to stay abreast of technical innovations in this rapidly advancing area.

On the other hand, imposing new burdensome regulatory obligations relating to the use of nucleotide sequence data would seriously hamper and chill international scientific collaborations that are mutually beneficial. We do hope that such a situation will never take place.

Joint statement by public and private sector organizations, academic and scientific institutions, data repositories and collections representing a broad range of stakeholders

SUBMISSION

Background

The public and private sector organisations, academic and scientific institutions, data repositories and collections underwriting this statement represent stakeholders who are actively engaged in or support the conservation and/or sustainable use of genetic resources to unlock their potential for society in different domains.

Most have also been active and collaborative participants in negotiations related to Access and Benefit Sharing (ABS) for many years, providing expertise and insights on the systems and practices which could best support effective ABS related to genetic resources.

As key stakeholders, the signatory organisations are vigilant about the potentially harmful effect of inappropriate or overly burdensome regulation of genetic resources. They are therefore greatly concerned about proposals to apply ABS obligations to DSI. Such obligations would place additional hurdles on

biological research – with potentially negative consequences for the advancement of science and the huge societal value this generates, as well as for achieving the three objectives of the CBD.

Statement of the issue

The unencumbered access to and use of DSI in the public domain benefits countries at all levels of development – it supports conservation, fosters research into technological solutions to tackle societal challenges, and benefits the population as a whole. Researchers collect and extensively use DSI to advance science and scientific understanding of biological systems. The rate of scientific advancement and technological development is heavily dependent on unencumbered access to and use of publicly available DSI. Barriers to the sharing and use of DSI would discourage innovation and scientific research. Extensive tracking and tracing mechanisms would be needed – if they were even possible – ultimately making downstream uses more complex and costly, and products and technologies less accessible. The net effect on conservation and sustainable use of biodiversity could be negative and in contradiction with the objectives of the CBD and the NP, as well as with several of the UN Sustainable Development Goals.⁴⁶

Recommendations

The signatory organisations emphasize the importance of continued ready access to, and use of, DSI for advancing research and development that will help achieve the objectives of the CBD and the NP, and recommend that CBD Parties:

1. Acknowledge the importance of unencumbered dissemination and use of DSI, including for fulfilling the objectives of the CBD and the NP.
2. Explicitly recognise that the open sharing of DSI represents a form of non-monetary benefit sharing.
3. Recognise the need for more fact-based information on how DSI is collected, generated, shared and used in biological research, as well as on the value which the unencumbered access and use of DSI represents for furthering the objectives of the CBD and the NP. Such value is created by, for instance, facilitating biological research, taxonomic studies to deepen the understanding of biodiversity, and international research collaboration. The submissions by a wide variety of stakeholders already provide examples in this regard.
4. Acknowledge the important role of major publicly accessible global databases holding DSI in providing an effective and inclusive system for sharing DSI globally.
5. Encourage capacity building to further enhance access to and use of DSI, in particular by developing countries and with a focus on their needs.
6. Recommend that international and national efforts should be focused on the effective and internationally coherent implementation of the current ABS framework to realise CBD objectives, instead of on reopening negotiations on the scope of the CBD.

Justification

The signatory organisations wish to emphasise that the outcome of the discussion on DSI will have a very important impact on the future of biological research and its benefits for society, and urge CBD Parties to take the following into consideration:

1. As noted above, DSI is a critical tool in the conservation and sustainable use of genetic resources, and its use is currently leading to many societal benefits. Impeding the flow and use of information would work against research projects in several different areas, including many which contribute to the objectives of the CBD and the NP. The CBD specifically encourages the exchange of information to support conservation, protection and sustainable use of biodiversity⁴⁷. It is of paramount

⁴⁶ Including, for example, goal 2 on hunger and food security, goal 3 on health, goal 13 on climate change and goal 15 on biodiversity, forests and Desertification

⁴⁷ Article 17, Convention on Biological Diversity

importance that DSI in the public domain continues to remain freely accessible to achieve the broader aims of the CBD.

2. Making the sharing of DSI more legally uncertain, time-consuming, administratively burdensome and costly by regulating its use will have a chilling effect on the use of DSI in biological research, and consequently lead to less benefits. All countries rely on and benefit from unencumbered access and exchange of DSI to deal with key societal issues, like human, animal and plant health, food security and the environment. There is currently a well-established and functioning international framework supporting the open exchange of DSI, consistent with established principles of ethical and responsible scientific research that are foundational for the advancement of science. Open exchange is especially important for DSI, whose unencumbered use allows the swift compilation, comparison and reanalysis of genetic information from a variety of sources, across multiple databases and gene sequences.
3. Considerable “non-monetary” benefits can be derived by all countries from the open exchange of publically available DSI. The extensive public DSI database managed by the INSDC (International Nucleotide Sequence Database Collaboration), maintained at the expense of the three host governments, is freely accessible to researchers in all countries; and was accessed by 172 countries from all regions between 2014 and 2016⁴⁸. Open exchange of DSI is also necessary for international research collaborations, which not only allow the pooling of expertise and resources to resolve problems of global or regional relevance, but are also essential vehicles for capacity building and the exchange of knowledge and expertise. Countries also benefit from the technologies and products resulting from research supported by open exchange of DSI, wherever this research takes place. Widespread regulation of DSI is likely to result in a dramatic reduction in information being made available in public databanks. There is therefore a danger that the significant non-monetary benefits currently being delivered to developing countries could be diminished in the process.
4. Academic and public institutions, which are key players in the R&D process, will be seriously affected by ABS obligations on DSI. Such obligations would also seriously impede international collaborations in which such institutions participate.
5. Under the existing NP framework, benefit-sharing provisions relating to DSI can already be included by provider countries in mutually agreed terms. In order to include DSI within the scope of CBD and/or NP, contracting Parties to the CBD and the NP would have to renegotiate the scope of the treaties. Numerous legal interpretations have confirmed that the definition of genetic resources refers to tangible material and does not include immaterial information. Negotiations to change this definition would require years, or even decades, as well as resources that would be better spent on efficient implementation of effective and workable measures that achieve the three objectives of the CBD.

League of European Research Universities (LERU)

SUBMISSION

Digital Sequence Information = Nucleotide Sequence Data! But more clarity is needed on its scope.

Introduction

⁴⁸ “As a member of INSDC (International Nucleotide Sequence Database Collaboration) maintaining the international DSI database, we found that a total of 172 countries were using the database during the period from 2014 to 2016 (Europe 46, Asia 43, South/Central America 35, Africa 32, Oceania 14, and North America 2; Website access: 1,621,300) upon investigating the state of the use from outside Japan of DDBJ (DNA Data Bank of Japan) operated by National Institute of Genetics”- *Extract from the Submission by the Government of Japan on the “Current state of the use of digital sequence information on genetic resources in the biodiversity field”*.

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 (the Nagoya Protocol on Access and Benefit Sharing (ABS)) is a supplementary agreement to the 1992 *Convention on Biological Diversity* and entered into force on October 12th, 2014. It aims to ensure a fair and equitable sharing of benefits arising from the utilisation of genetic resources and, if relevant, the use of any associated traditional knowledge (aTK).

While accepting the benefits of the use of Digital Sequence Information (DSI) for the three objectives of the Convention on Biological Diversity, the League of European Research Universities (LERU) strongly disagrees with the proposal to include DSI under the Nagoya Protocol. Inclusion of DSI would be against the spirit of open science, would damage environmental, life sciences and biodiversity research, and would be detrimental to less developed countries. Open access to DSI allows scientists from all over the globe - irrespective of finance, status or location - to be able to do molecular research. This undoubtedly aids provider countries as well as preventing single users from profiteering from what they have sequenced. Sharing data leads to cost savings in research projects and adds value to the data. Without open access to DSI, such research will be significantly hindered.

There is considerable uncertainty over the term DSI. LERU has submitted this short note to the Convention on Biological Diversity to outline both what we believe is meant by the term DSI and what DSI covers, and to outline some questions which need clarifying to ensure that the term is both legally robust and workable⁴⁹.

The Need for a Globally Accepted Legal Definition of Digital Sequence Information

There is currently no universally agreed definition of the term DSI⁵⁰. However, a clear universally agreed definition of DSI is essential in order to assess the potential impact of including DSI within the scope of the Nagoya Protocol. Hence, before discussing the potential for including DSI within the Nagoya Protocol, the scope of this discussion should be clearly focused, based on a shared, accepted and scientific definition of DSI.

It is our firm belief that DSI, for the overwhelming majority of scientists, only encompasses nucleotide/genetic sequence data, nucleotide sequence information or genetic sequences, which we term ‘nucleotide sequence data’. These are currently found within open access databases such as the INSDC (International Nucleotide Sequence Database Collaboration). Occasionally, it may be extended to encompass protein sequence data as well. It would be our preference to use the unambiguous term ‘nucleotide sequence data’ in place of the current placeholder term, ‘DSI’⁵¹ which is unclear and open to considerable interpretation.

If DSI is considered within the scope of the Nagoya Protocol, it is crucial that there is a precise, and legally-binding definition of DSI. Without this, researchers, and other users, will risk inadvertently contravening the Protocol. Thus, LERU believes it is necessary to clarify what types of genetic sequences would be included under the umbrella term DSI. For example:

- Does DSI include genes, chromosomes, chromosome fragments, coding sequences, all of them or more?
- How should laboratory generated mutants (either developed through mass mutagenesis or precision approaches such as CRISPR) be handled? In LERU’s view, the researcher developing the mutant should also be credited for his/her part in generating the material.

⁴⁹ Part of this note was first published in November 2018, prior to the 14th meeting of the Conference of Parties to the Convention on Biological Diversity and the 3rd meeting of the Conference of the Parties serving as the meeting of the Parties to the Nagoya Protocol on Access and Benefits Sharing in Sharm-El-Sheik, Egypt, in November 2018. <https://www.leru.org/files/LERU-Nagoya-Statement-November-2018.pdf>

⁵⁰ Fact-Finding and Scoping Study on Digital Sequence Information on Genetic Resources in the Context of the Convention on Biological Diversity and the Nagoya Protocol (CBD/DSI/AHTEG/2018/1/3) <https://www.cbd.int/doc/c/b39f/4faf/7668900e8539215e7c7710fe/dsi-ahteg-2018-01-03-en.pdf>

⁵¹ For the rest of this paper, we use the term DSI, as this is the term currently used by the CBD.

- How should identical genetic sequences found across different countries be dealt with? For example, sequences from river waters, migratory species, etc.?
- What are the tolerance levels for variations within a DNA sequence? Different organisms exhibit very different rates of mutation.
- How should environmental sequence data be handled? Much of the sequence data recovered in this case may not be attributed to a specific organism.
- How are artificial sequences and those consisting of novel base pairs treated⁵².
- What would come under ‘scope’ – the specific use of DSI or the sequence?

In conclusion, LERU strongly supports further consultation with relevant parties in this area to develop a consensus view on what DSI is (and what it is not). Universities can, and should, play a major role in these discussions to ensure that the definition is fit for purpose and leaves no room for misinterpretation.

Perspective

The university sector carries out a large amount of sustainability and biodiversity research worldwide, mostly without commercial intent. Such research is vital at a time in which there is a growing acceptance of the urgent need for immediate, and more substantial, actions to promote conservation and prevent further biodiversity loss across the world.

LERU strongly believes that DSI should *not* fall within the scope of the Nagoya Protocol. In LERU’s view, the benefits for biodiversity and conservation research arising from free and unhindered open access to digital sequence information far outweigh any potential (financial) benefits which providers could gain from restricting access to DSI using Nagoya Protocol frameworks.

LERU believes there are still too many unknowns over what DSI is, and how DSI could be incorporated into the Nagoya Protocol, to make it a viable option at present. LERU is delighted that the Convention on Biological Diversity is actively seeking opinions on the DSI concept, which we hope will result in a robust description for DSI. LERU believes the term ‘nucleotide sequence data’ is a perfect and precise term to use in place of DSI. Only when there is an agreed definition of DSI (or another appropriate term) can any system for ABS be investigated. The university sector, as a key provider of knowledge and capacity sharing on biodiversity, should be fully consulted at every stage to ensure that a robust and workable system is developed.

Natural History Museum UK, Royal Botanic Gardens Kew, and Royal Botanic Gardens Edinburgh

SUBMISSION

A submission from:

The Natural History Museum, Cromwell Road, London SW& 5BD, UK
Royal Botanic Gardens Kew, Richmond, Surrey TW9 3AE, UK
Royal Botanic Garden Edinburgh, 20A Inverleith Row, Edinburgh, EH3 5LR

In Response to notification 2019-012 on 5 February 2019:

The Conference of the Parties to the Convention on Biological Diversity, at its 14th meeting, in decision 14/20 requests Parties, other Governments, indigenous peoples and local communities, relevant stakeholders and organizations to submit their views and information on Digital Sequencing Information⁵³:

⁵² <https://phys.org/news/2017-10-letters-dna-functions.html>

⁵³ <https://www.cbd.int/doc/decisions/cop-14/cop-14-dec-20-en.pdf>

- (a) To clarify the concept, including relevant terminology and scope, of digital sequence information on genetic resources and if and how domestic measures on access and benefit-sharing consider digital sequence information on genetic resources;
- (b) On benefit-sharing arrangements from commercial and non-commercial use of digital sequence information on genetic resources.

Executive Summary

We recommend that term ‘Nucleotide Sequence Data’ (NSD) is used instead of “DSI”, and that this concept be taken as the basis for discussion. NSD is the DNA (or RNA) sequence - the arrangement of nucleotides⁵⁴ in a molecule of DNA or RNA. Our reasons are:

- i. A genetic resource when accessed is not accompanied by information about properties or other products of research. The concept of “DSI” closest to an accessed tangible genetic resource is the nucleotide sequence, without its properties or other information developed through research. Neither nucleotide sequence data nor a tangible genetic resource have intrinsic information but can only be considered as providing data;
- ii. Use of “data” avoids use of the term “information”, since information, as the result of research on a genetic resource, might attract intellectual property rights which would not a priori fall under providing country sovereign rights;
- iii. Inclusion of the term ‘Digital’ might be taken to limit the applicability of any discussion to content available digitally (i.e. in computer databases or online) only;
- iv. The term ‘Digital Sequence Information’ is not used in the taxonomic, systematic and conservation communities and not understood

Non-commercial scientific research, including that using NSD, increasingly uses an open access model. Funders and publishers demand that data are published and made openly available, and the public databases that hold the overwhelming majority of sequence data have an open access policy requirement. This model supports scientific research and environmental management globally; users of these databases are in every Party to the CBD. The open access model supports scientists in all countries, noting that research results become more reliable with increasing amount of data on which they are based. Open release of NSD for use by the global scientific community is clearly an example of benefit-sharing.

We recognise that a lack of capacity in many countries hinders greater use of NSD. Non commercial biodiversity research organisations are engaged in capacity building to meet this problem through training at graduate, postgraduate and professional levels, and by supporting infrastructural growth in provider countries. This is considered good practice within the sector.

Benefits derived from NSD use are mostly applicable on a multilateral basis; many projects make use of NSD generated from multiple geographic sources and publish on taxa that occur in more than one country. Unrestricted publication of data and literature is the most effective model for enabling this.

Our view is that the value of openly publishing NSD and thus making non-monetary benefits derived through their use⁵⁵ available to users in all countries outweigh the risks of these data being used for commercial purposes without monetary benefits being shared with countries from which the original genetic resources were accessed. However, more research is needed to identify cost-effective solutions that mitigate the risks without increasing the risk of loss of non-monetary benefits.

Increasingly countries are including some form of “DSI” in their domestic ABS legislation. The noncommercial sector ultimately needs clarity on what is and is not covered by their usages. This issue

⁵⁴ Nucleotides are subunits that are polymerized (connected to form a linear chain) to make nucleic acids (DNA and RNA). The four types of nucleotides in DNA and RNA are Adenine, Thymine, Guanine, and Cytosine, which are usually abbreviated to A, T, G and C. In RNA Thymine is replaced by Uracil (U). The order in which these nucleotides occur in a strand of DNA is the DNA (genetic) sequence.

⁵⁵ See the 2017 submission from NHM, RBG Kew and RBG Edinburgh for a discussion of these benefits

cannot be ignored, since some national legislation and permitting procedures are beginning to restrict access to and use of genetic resources through these bilateral agreements, and consequently reducing the development of non-monetary benefits that can be shared.

Preamble

The three institutions submitting this document are committed to supporting implementation of the CBD and to benefit-sharing in our day-to-day operations. Our operations are primarily noncommercial, and our utilisation of genetic resources does not generate monetary benefits. Our benefit-sharing is consequently predominantly in terms of non-monetary benefits. For DSI as discussed below we believe that the most effective basis for benefit-sharing is on a global basis in the manner required by Aichi Target 19. This is already a model that is operating, as discussed below.

A summary of our previous submission on “Digital Sequence Information” and its importance to the CBD, made in response to notification 2017-37 issued on 25 April 2017⁵⁶, is provided below:

- The generation of “DSI” is growing rapidly in terms of the number of individual organisms and species being sequenced and the depth of genomic coverage obtained per sample.
- Three global databases, known as the International Nucleotide Sequence Database Collaboration (INSDC), exchange information and currently mediate data on more than 200 million sequences; one of them (EMBL-European Bioinformatics Institute (EMBL-EBI)), is subject to more than 100 million searches a year.
- Expecting the large-scale open-access international databases to regulate the use of ‘DSI’ from them is impractical and inappropriate: (i) the publisher cannot be responsible for compliance with any conditions by its users; (ii) the INSDC databases are required by their core policies (approved by national / regional authorities) to avoid both barriers to downloading the sequence data and applying conditions on their use.
- Publicly-available ‘DSI’ provides a vital benefit for biodiversity conservation and sustainable management of natural resources and supports the implementation of the Convention. The INSDC databases form *de facto* a part of the Global Taxonomic Information System called for in COP decisions, and the sharing of knowledge and the science base required for the achievement of Aichi Biodiversity Target 19.
- No country holds sequence data for all of its biota and species likely to be intercepted by quarantine as Alien Species, pests etc. This would still be the case if sequence data were treated as a bilateral benefit between researchers and provider country. The only way in which Parties can obtain sequence data for supporting implementation of the Convention is through the freely-available global databases.
- Our unequivocal view is that sharing “DSI” without hindrance is overwhelmingly beneficial to Parties to the Convention, assisting them to understand and manage their biodiversity through, *inter alia*, establishing identities and names to taxa, understanding genetic structure of populations for effective conservation management, assessing the diversity of organisms present in a given site, detection of protected or invasive species, identification of species in trade, understanding migrations, effects of harvest rates on genetic diversity, understanding changes in genetic diversity over time, phylogenetic diversity across landscapes to assist large scale conservation planning, pollinator conservation.
- Any modification of the current model of use of “DSI” would risk limiting the non-monetary benefits indicated above, and consequently the implementation of the Convention. The financial equivalence of these benefits has not been assessed, but before any action is taken it would be helpful to make this calculation and compare it (plus the implementation costs) to the revenues that might be generated by alternative models. Monetary benefits may also not be realised, as legal uncertainty and audit complexity would likely discourage usage of biological materials.

⁵⁶ <https://www.cbd.int/abs/DSI-views/NHM%20RBGK%20RBGE-DSI.pdf>

We argue in this current document that “DSI” be replaced by the more precise term “Nucleotide Sequence Data” (NSD). Our previous comments using the term “DSI” apply to NSD.

On the concept, including relevant terminology and scope, of digital sequence information

The term ‘Digital Sequence Information’ is not used by taxonomic, systematic and conservation research and we find it unhelpful. There is no agreed meaning, and we are concerned that it, and similarly imprecise terminology, is beginning to be used in bilateral contracts (permits, ABS contracts etc) and national legislation, consequently introducing legal uncertainty. This is causing reluctance by scientists to carry out research in countries with such legislation. We explore below what the term means to us and propose an explicit meaning and possible alternative terminologies.

A tool now vital for taxonomy, systematics and conservation globally is the analysis of the arrangement of nucleotides on a strand of naturally occurring DNA or RNA – the ‘Nucleotide sequence’⁵⁷. These are ‘Nucleotide Sequence Data’ (NSD). NSD include non-coding & coding sequences, regulatory sequences, conserved sequences, genes that encode specific traits, DNA without known function and ‘junk DNA’. Larger data elements would include the entire genome of an organism [or, indeed, of a clade (pangenome) or environmental sample (metagenome)]. NSD are aggregated from naturally occurring genetic resources generated as a part of research or downloaded from INSDC⁵⁸ and other databases. Analyses of NSD are interpreted in research to develop understanding of biological diversity at genetic, species and ecosystem levels.

We distinguish between ‘data’, which are observations of naturally occurring conditions lacking extrapolated meaning, and ‘information’, which emerges through cognitive (or other) processing and application of data - research. Since *information* is developed through processing or analysis of data it can potentially be covered by Intellectual Property Rights of the researcher and therefore a priori cannot come under sovereign rights of a country from which the data were sourced⁵⁹. This is not to exclude the potential for benefit-sharing with that country⁶⁰. Indeed, our practice is to share such information freely and we understand this as part of our benefit sharing activities. Consequently, we regard NSD as intrinsic factual data directly derived from the genetic resource. Any interpretation or annotation of these facts are here regarded as information, and as benefits that can be shared.

When a genetic resource is accessed *in situ* in a provider country, all that is acquired is an organism or set of organisms. These lack a known identity, and there is no attached information; any information has to be provided through research or pre-existing knowledge of the researcher. Discussion of “DSI” in the context of sovereign rights would logically be an analogue of what is accessed as a material entity, focusing on the genetic resource element of that entity. For this reason we propose that the concept of “DSI” be restricted to NSD.

Our definition of “DSI” is therefore the arrangement of nucleotides on a strand of naturally occurring DNA or RNA, and whenever we refer to “DSI” this is what we intend. We note that this clear concept is in line with that of ‘Genetic Sequence Data’ (GSD) proposed by some Parties to the CBD. We propose that either ‘GSD’ be adopted as a working term, or the more precise ‘Nucleotide Sequence Data’ (NSD). We specifically exclude ‘digital’ from our proposed terminology to avoid inappropriate restriction to a single condition of storage or transmission medium of aggregated data from genetic resources.

Benefit sharing from non-commercial use of digital sequence information on genetic resources

Bilateral benefit-sharing arrangements are general practice when research is carried out on Genetic Resources accessed with PIC and MAT under the ABS requirements of many countries. Where NSD are generated through utilisation of Genetic Resources under such a traditional bilateral agreement and used

⁵⁷ See footnote 47 above

⁵⁸ International Nucleotide Sequence Database Consortium

⁵⁹ The Annex to the Nagoya Protocol specifically lists joint ownership of IPR in its list of non-monetary benefits, acknowledging that this is a benefit developed by research.

⁶⁰ See World Intellectual Property Organisation (WIPO), 2018, A guide to intellectual property issues in Access and Benefit-sharing Agreements. WIPO, Geneva.

subsequently in the same research then bilateral benefit-sharing arrangements may be in place. However, we have no experience of bilateral benefit sharing as a result of using exclusively publicly-available NSD.

In addition to bilateral benefit-sharing, the results of non-commercial research in our sector, and the data that underly those results, are made public, increasingly on an open access basis. Thus scientific data, information and research results are made globally available. Scientific research has always operated on a multilateral benefit-sharing model (although not termed as such). No Party to the Convention has the capacity to generate or manage data or information on all of its biota. Instead all must rely on data and information generated and held elsewhere. If scientific data and information were treated solely in a bilateral benefit-sharing manner, countries would not benefit from information generated about non-endemic species, or from *ex situ* collections. Data sharing is particularly applicable to NSD which, once used for research (and sometimes before research is published), are made publicly available through open access databases. In our previous submission we discussed these databases and noted that their development and accessibility was de facto a contribution to the global taxonomic information system as called for in CBD COP IX/22. We also showed that it is a response in part to CBD Parties' calls for availability of information in a number of other COP decisions, including COP IX/22, COP X/39, COP XI/29 and the Aichi Biodiversity Targets, particularly Target 19. We reiterate that scientific research and use of NSD analysis for environmental management, is increasingly effective with the increased quantities of available data. Global open-access systems are the most effective way of achieving this.

Publishing NSD held privately within our institutions is a form of benefit sharing, as well as a requirement of modern science. The uses to which these data are put in support of countries' activities to conserve and sustainably use their biodiversity were discussed in our previous submission. That NSD and other related data are being used globally can be seen from the EMBL-EBI's live data map⁶¹ which shows live use of their databases; a recent analysis showed that users every country on Earth made use of the INSDC databases⁶².

We are aware that there are capacity needs in some developing countries that limit use of NSD, although we are also aware of researchers in many countries making effective use of these data. We view capacity building in making use of NSD as a part of benefit sharing. We engage in this in several ways:

- Our institutions all run training courses including international students. These courses develop skills in DNA sequencing and analysing nucleotide data. This raises capacity in the students' countries to make use of NSD.
- Scientists in our institutions all support and nurture PhD students, including from developing countries, and some of these at least are engaged in the degeneration and analysis of NSD.
- Our institutions play host to research partners and visitors who make use of our molecular laboratories to carry out sequencing.
- Our laboratories regularly carry out sequencing tasks at cost for external researchers, both of tissues they submit and of tissues held in our own collections.
- Many of our researchers are in research partnerships with researchers in other countries, including developing countries, and will generate NDS as a part of their joint research. In this way, where sequencing facilities are unavailable elsewhere, we can provide that capacity at a far lower cost than would otherwise be required.

While the publication of NSD developed in our research is, we submit, a benefit, and one supplemented by our capacity building, the research results we publish arising from the use of published NSD also provide benefits. We discussed the use of such publications in the implementation of the CBD in our previous submission. As noted, publications from our researchers are made available globally. Increasingly we are seeking to publish in open access journals so that payment is not an issue. Because we work in a global community of research papers tend to be shared with our peers in both developing and developed countries,

⁶¹ <https://www.ebi.ac.uk/web/livemap/>

⁶² Leibnitz Association, 2018, The DSI debate: a primer on the science and infrastructure behind DSI. Discussion paper.

ensuring that research results are seen rapidly by scientists most able to use them. Such sharing of results is not documented in any standard way, but is a regular part of our activities.

A direct way in which we share benefits of NSD is in our use of it for identification. Every year, our staff identify numerous organisms for bodies in developing countries, often through the use of online data for comparison. Such identifications might not be possible without this resource.

Risks

We have presented evidence that use of NSD contributes to knowledge of biodiversity, both academic and applied in the context of the CBD. We are aware that there is concern among some countries that benefits (particularly monetary benefits) arising from the use of NSD are not being shared. However, we are unaware of any analysis of the actual value and development costs of commercial products generated following use of NSD. We are concerned that attempts to manage use of NSD will have a negative impact on non-commercial use and consequently on biodiversity research and environmental management.

Restrictions applied in some access permits, for example prohibiting the dissemination of sequence data to third parties, or extracting sequence data being only allowed by the GR collector are already having a negative impact. Researchers are unable to carry out their research effectively, and restriction on publishing means that biota from the country are excluded from the research. In addition to these constraints, some countries are explicitly stating their sovereign rights over “DSI” and bringing it into their domestic ABS legislation. As yet there is no clarity on how such legislation can be applied in practice, on the meaning of “DSI” or the legal implications of such legislation in other countries. Without some progress in international agreements we see problems in carrying out research, and sharing benefits as required by the CBD, increasing to the detriment not only to research but to biodiversity conservation and sustainable use.

<p>The Natural History Museum Vienna (NHMW)</p>
--

SUBMISSION

The Natural History Museum Vienna (NHMW)

- 1) being one of the largest institutions of its kind and one of the most important biodiversity research facilities in the world,
- 2) conducting non-commercial, publically funded fundamental research on the distribution, evolution, origin and extinction of biodiversity at different levels (molecular, morphological, ecological), with a major part of that research based on its scientific collections,
- 3) being engaged in non-monetary benefit sharing and capacity building by international training and collaboration
- 4) having carefully read the Note by the Executive Secretary of the CBD’s Subsidiary Body On Scientific, Technical and Technological Advice
- 4) considering the multiple aspects covered by the text of the Note and the Annex
- 5) supporting CETAFs proposal to replace the term ‘Digital Sequence Information (DSI)’ by ‘Nucleotide Sequence Data (NSD)’
- 6) being concerned that development of restrictions on the use of NSD will create serious impediments to fundamental biodiversity research and achievement of the major CBD goals

provides the following statement:

At the NHMW (like in many other museums and research institutions worldwide), DNA is frequently extracted, and corresponding sequences are generated from collection material for taxonomic, phylogenetic, or evolutionary studies, that are published and thus shared with the global research community. The main mission of natural history collections is to contribute to a better understanding and thus protection of

biodiversity. These aims are non-commercial and frequently pursued in the frame of international co-operations. We strongly discourage making any decision further impeding this mission. While our major points are outlined above, as a member to CETAF, we fully support the view expressed by the statement submitted by the consortiums secretary and add some general and more detailed remarks to the following three points:

1. The concept, including relevant terminology and scope, of “digital sequence information” on genetic resources
2. Domestic measures on access and benefit-sharing considering “digital sequence information” on genetic resources;
3. Benefit-sharing arrangements from non-commercial use of “digital sequence information” on genetic resources.

Comments in detail

1. The concept, including relevant terminology and scope, of “digital sequence information” on genetic resources

As outlined in detail in the CETAF statement, the term “Digital Sequence Information” should be replaced by “Nucleotide Sequence Data” (NSD) to reduce the inherent terminological ambiguity to a minimum and give more legal security.

2. Domestic measures on access and benefit-sharing considering “digital sequence information” on genetic resources

NSD, specifically when acquired in biodiversity research, is extremely important for decisions on conservation of biological diversity and sustainable use of its components. Thus the aims of the CBD rely on free access to and use of this data. Following good scientific practice, NSD is made publicly available in online databanks (NCBI GenBank, BOLD, Dryad etc.), which is also required from authors by most scientific journals. We follow this good scientific practice with all above-mentioned sequences generated at the NHMW. Open access to such information not only warrants reproducibility but also provides this data (generated under standardized quality norms) to the global research community so that all benefit from it. The hosting of NSD represents a global benefit, and these portals contribute significantly to a sustainable future. We consider it a serious threat to these benefits, should bilateral agreements prior to using NSD become mandatory.

3. Benefit-sharing arrangements from non-commercial use of “digital sequence information” on genetic resources.

The NHM already contributes to non-monetary benefit sharing. The objects kept in our collections are not only used for our own research, we also provide material and training to researchers worldwide and are embedded in a network of international collaborations including partners from both developed and underdeveloped countries. Further, we provide access by hosting thousands of visitors per year, among others within the SYNTHESYS+ framework. Among the shared benefits arising from these activities are training, transfer of knowledge, and joint publications. Evidently, we, as part of the global research community, contribute considerably to the knowledge on global biodiversity, usually on public funding.

To summarize, open access of NSD in our view fully supports the goals of the CBD. Since mostly generated by publicly funded projects, open access to NSD may also fall under Freedom of Information (FOI). Further, the regulation of NSD by bilateral agreements under the Nagoya Protocol would severely counteract the aims of the CBD, hampering fundamental research on biodiversity and limiting the share of common benefits (in the sense of common knowledge) generated by global biodiversity research.

Society for the Preservation of Natural History Collections (SPNHC)
--

SUBMISSION**Preamble**

At the 2018 Conference of the Parties to the CBD and Meeting of the Parties to the Nagoya Protocol in Egypt decisions were adopted to gather information on Digital Sequence Information (DSI) in order to inform further discussion by the Parties. The decision requests Parties, other Governments, indigenous peoples and local communities, relevant stakeholders and organizations to submit their views and information:

- (a) To clarify the concept, including relevant terminology and scope, of digital sequence information on genetic resources and if and how domestic measures on access and benefit-sharing consider digital sequence information on genetic resources;
- (b) On benefit-sharing arrangements from commercial and non-commercial use of digital sequence information on genetic resources.

The areas of research that use “DSI” are characterised by rapid growth and innovation. The research carried out by the SPNHC membership in the realms of taxonomy, systematics, description, environmental and biodiversity studies, evolutionary development, conservation support etc. are no different in this respect. Terminology applied to “DSI” and modalities of its use when agreed need to be unambiguous and ‘future-proofed’ to whatever extent possible. This will provide certainty and a firm base to carry on research and benefit sharing.

This consideration will affect the characterisation of “DSI”, and results in our proposal below.

Currently, the vast majority of research results from all countries are made publicly available, as are the data used in scientific studies, which includes “DSI”. The ability to replicate other researchers’ results is a cornerstone of all scientific research, and this requires the data to be available. That “DSI” is freely available has raised concerns among some developing countries that it could be used in commercial applications without triggering an obligation to share benefits with the provider countries from which the original genetic resources were acquired. Our members understand this concern (although this does not apply to the vast majority of research carried out by the SPNHC membership, which is non-commercial), but also wish to highlight the large number of non-monetary benefits which are shared by the scientific community which actively contribute to and support the objectives of the CBD. These benefits were discussed in our submission to the Executive Secretary in 2017⁶³ and will not be repeated here.

Research data including “DSI”, when published, are maintained to the standardised quality norms of the global research community and available for use in Provider and User Countries at zero marginal cost. This important function of science is reflected in Article 15 of the CBD, which calls Parties to take *legislative, administrative or policy measures ... with the aim of sharing in a fair and equitable way the results of research and development and the benefits arising from the commercial and other utilization of genetic resources with the Contracting Party providing such resources.*

Our members are committed to benefit-sharing and SPNHC as compiled information on best practices and organised workshops for its members to facilitate this⁶⁴. For “DSI”, the Legislation and Regulations Committee of SPNHC believes that the most effective basis for benefit-sharing is on a global basis for the common good, in the manner required by Aichi Target 19. This will require a functional common set of technical standards. Promising practical examples demonstrating such standards are available, including the International Nucleotide Sequence Data Collaboration (INSDC), Global Biodiversity Information Facility (GBIF) and the Barcode of Life Database (BOLD), which address both technical and legal issues, in the case of GBIF explicitly operating within an intellectual property rights framework. We think building

⁶³ <https://www.cbd.int/abs/DSI-views/SPNHC-DSI.pdf>

⁶⁴ [https://spnhc.biowikifarm.net/wiki/Access_and_Benefit-Sharing_\(Nagoya_Protocol_and_the_CBD\)](https://spnhc.biowikifarm.net/wiki/Access_and_Benefit-Sharing_(Nagoya_Protocol_and_the_CBD))

on these should be preferred instead of implementing new systems that aim to restrict and regulate “DSI”. Because of the exponentially growing amount of data, developing new systems with additional legal and policy-related requirements would be a difficult and expensive task with unknown results but potentially negative impact on science globally, and particularly on CBD implementation.

This submission builds on the 2017 submissions from SPNHC⁶⁵ and CETAF⁶⁶ to the Executive Secretary on Digital Sequence Information on genetic resources – benefits of their use and their public availability for the three objectives of the Convention on Biological Diversity, and ramifications of restricting access to DSI.

Contents

Comments in Detail

1. The concept, including relevant terminology and scope, of “digital sequence information” on genetic resources
2. Domestic measures on access and benefit-sharing considering “digital sequence information” on genetic resources;
3. Benefit-sharing arrangements from non-commercial use of “digital sequence information” on genetic resources.

Summary

Comments in detail

1. The concept, including relevant terminology and scope, of “digital sequence information” on genetic resources

The term “Digital Sequence Information” is not used by the SPNHC membership in their work and appears to be limited to policy discussions⁶⁷. The very diverging interpretations of the term now current and the resultant ambiguity make it important to use a different term of precise meaning.

One issue lies in the term ‘information’. We understand that while ‘data’ are observations of naturally occurring states lacking extrapolated meaning, ‘information’ emerges through cognitive (or other) processing and application of data. In the “DSI” context ‘data’ are extracted from naturally-occurring genetic resources – the genetic resource itself when accessed has no intrinsic ‘information’. The relevant data are the arrangement of nucleotides on strands of naturally occurring DNA or RNA⁶⁸. This is ‘Nucleotide Sequence Data’ (NSD). Information about the genetic resource arises through analysis of these data.

Free sharing of both data and information, enabling researchers and other users in Providing and other Countries is the usual practice in the non-commercial sector, and we view such provision as constituting a shared benefit. However, there are important differences between the two.

In practice, since ‘information’ is developed through analysis of ‘data’, it is potentially covered by Intellectual Property Rights (IPR). Sharing of IPR is identified as an example of a non-monetary benefit in the Nagoya Protocol Annex; consequently, it can be seen that IPR are a result of research and not under sovereign rights of a country as are natural resources (genetic resources).

⁶⁵ <https://www.cbd.int/abs/DSI-views/SPNHC-DSI.pdf>

⁶⁶ <https://www.cbd.int/abs/DSI-views/CETAF-DSI.pdf>

⁶⁷ Laird & Wynberg, 2018, A Fact-Finding and Scoping Study on Digital Sequence Information on Genetic Resources in the Context of the Convention on Biological Diversity and the Nagoya Protocol.

CBD/DSI/AHTEG/2018/1/3

⁶⁸ Nucleotides are the subunits that are connected into long chains to make nucleic acids (DNA and RNA). The four types of nucleotides in DNA are Adenine, Thymine, Guanine, and Cytosine, and in RNA Thymine is replaced by Uracil. The five nucleotides are usually abbreviated to A, T, G, C and U. The order in which these nucleotides occur in a strand of DNA or RNA is the DNA or RNA sequence or Nucleotide Sequence.

Analysis of NSD is a key tool for taxonomy. The data used for analysis are aggregated from naturally occurring GR and downloaded from databases such as GenBank and the other INSDC⁶⁹ members' databases, and include non-coding and coding sequences, regulatory sequences, conserved sequences, genes that encode specific traits, and 'junk' DNA⁷⁰; sequences do not need to have a known function to be of relevance in our research. There is no maximum size for a usable sequence. Analysis might be of single genes, multiple genes, entire genomes of organisms, of a clade (pangenome) or environmental samples (metagenomes). The results of analysis are interpreted to further our understanding of biological diversity.

We propose that discussion on "DSI" distinguishes between data (NSD), and information. We also propose that the concept of "DSI" be restricted to NSD. We note that this clear concept is also in line with 'Genetic Sequence Data' (GSD) as proposed by some Parties to the CBD. We specifically exclude 'digital' from our proposed terminology to avoid inappropriate restriction to a single current means of data storage and transmission of aggregated data from GR.

For the vast majority of taxonomic and systematic research, it is not the discovery and application of functions of the genes *per se* that are important. In any case, since the function can only be discovered by experiment or through the application of existing knowledge to predict or test for function (i.e. extrapolation of data including through automatic means) this would come under the heading of information.

Sequence data may be associated with a set of other data to increase its scientific value, such as:

- i) Collection site of the organism or sample from which the NSD was obtained;
- ii) the date on which it was collected;
- iii) the name of the collector;
- iv) the place where a physical voucher is stored (if it is retained) and the unique identifier of that voucher;
- v) the taxonomic name of the organism from which the DNA was sequenced.

While we do not consider this to be NSD, it is helpful contextual information and, where appropriate (and where it exists) it can be made available with NSD to which it applies. Associating these data is scientific best practice, but far from all sequences stored in public databases are associated with all of these data. Permit conditions may be stored as part of the record.

2. Domestic measures on access and benefit-sharing considering digital sequence information on genetic resources

We are aware that some countries are introducing domestic legislation including NSD. This gives us cause for concern. One is terminology; countries are using a variety of terms that are sometimes of unclear meaning, are often used inconsistently and thus increase legal uncertainty, instead of reducing it. Another is practical. Our members expect to seek permits or their equivalent when collecting biological specimens for research or addition to collections, and to share agreed benefits. However, these are negotiated and set out in bilateral arrangements, made in the context of significant investment of time and resources to establish scientific work in one country. Use of NSD includes work that is very different. Often the research is not targeted at a particular country but at a particular species or group of species (perhaps hundreds or even thousands of species). The uses to which the NSD are put include for example simple comparison to identify an organism or taxon from which a similar sequence has been isolated (or such a sequence within an eDNA sample, for example), developing hypotheses of evolutionary relationships and assessment of biodiversity richness. In the overwhelming majority of cases in our sector the outputs are information-based and shared globally. If countries require bilateral agreements before NSD are used this is likely to have several detrimental outcomes. Some researchers may simply ignore the requirement, exacerbating mistrust and leading to breakdown of both relationships and the growing incorporation of ABS in research workflows globally. Others will develop algorithms to avoid use of NSD from countries making such requirements, leading to fewer data and less information being generated to address biodiversity

⁶⁹ International Nucleotide Sequence Database Consortium

⁷⁰ Sequences currently of no known function

management priorities. Such a limitation would compromise achievement of the Aichi Targets as well as national NBSAP goals. The reasons for these actions is that the time required to reach an ABS agreement is likely to damage research significantly; such delays for access to physical specimens are already problematic, and there is no reason to expect that the very much higher rate of use of NSD will be handled any more expeditiously. To put this into context, every weekday, well over 38 million requests are made to EMBL-EBI websites⁷¹.

3. Benefit-sharing arrangements from non-commercial use of digital sequence information on genetic resources.

In the statements of SPNHC⁷² and CETAF⁷³ submitted 2017 the value of “DSI” to the implementation of the CBD was discussed, and would withdraw nothing from that statement, which was based almost entirely on the concept of NSD discussed here. Implementation of the CBD is of course on a national basis, and thus countries implementing the CBD are making use of benefits developed through the use and generation of NSD.

SPNHC members are aware that they have to agree bilateral benefit-sharing arrangements as a normal part of Mutually Agreed Terms when accessing genetic resources. Insofar as the benefits are data or information (as opposed to capacity-building, for example), practice is to both share them bilaterally and to publish them so they are available globally. While we have many examples of bilateral benefit-sharing in the context of tangible genetic resources, we have no experience of benefit-sharing on a bilateral contractual basis exclusively for NSD.

Although contract-based bilateral benefit-sharing for NSD is not in our experience, global benefit sharing from analysis of NSD is a common experience and part of daily work. As noted above, the prevailing model of non-commercial scientific work is of open sharing of data and outputs. As a concomitant to publication of research results researchers are required to make NSD from their own databases publicly available once research is published, generally through the large public databases of the INSDC. Thus benefit sharing when these benefits are new genetic sequences is done on a global (multilateral) basis rather than a bilateral basis. This methodology is far more efficient and valuable to all users, since it allows access to sequences relating to species outside national borders, important for identification of invasive species, for example⁷⁴.

The global resource of NSD provided by the INSDC is in constant use worldwide; although the majority of users come from North America, Europe and China the data made available are used by every Party to the CBD⁷⁵ (see <https://www.ebi.ac.uk/about/our-impact> for a real-time visualisation of use of EMBL databases). While researchers can (and sometimes do) send direct links to uploaded NSD to provider countries, there is no evidence that this makes any difference to availability and use of the data developed and shared.

NSD are in use globally, but there are still capacity building needs to increase Parties’ ability to realise the benefits and exploit these data. The SCBD has supported training in DNA barcoding, which includes making use of the NSD in the BOLD system. INSDC members also offer training and a range of training materials. The SPNHC membership is also active in capacity building. This may take the form of training as a part of research, for example training students while working in labs in providing countries, joint research involving generation and analysis of NSD, in-house training at bachelor’s, master’s and PhD levels, and informally through professional contact. Our organisation includes 686 members from 24

⁷¹ <https://www.ebi.ac.uk/about/our-impact>

⁷² <https://www.cbd.int/abs/DSI-views/SPNHC-DSI.pdf>

⁷³ <https://www.cbd.int/abs/DSI-views/CETAF-DSI.pdf>

⁷⁴ Lack of such sequences has been identified as a problem for invasive Alien Species detection – see Lyal & Miller, 2018, Capacity of United States federal government and its partners to rapidly and accurately report the identity (taxonomy) of non-native organisms intercepted in early detection programs. 22pp. https://www.doi.gov/sites/doi.gov/files/uploads/lyal_federal_capacity_taxonomy_contractorreport_22october2018.pdf

⁷⁵ Leibnitz Association, 2018, The DSI debate: a primer on the science and infrastructure behind DSI. Discussion paper.

countries, including Africa, Asia, Australia, Europe, and North America and shares relevant information on Nagoya and ABS with its membership. Individual SPNHC members are actively engaged in capacity building with visiting researchers and colleagues from developing.

Summary

The term “Digital Sequence Information” is ambiguous and “DSI” is increasingly used as a convenient acronym without consideration of what it encompasses. We suggest that a replacement term be used in discussion and negotiation – ‘Nucleotide Sequence Data’ (NSD). This is the order in which nucleotides (Adenine, Thymine or Uracil, Guanine, and Cytosine) occur in a strand of DNA or RNA. The definition excludes ‘information’ which is developed through analysis of the data and which might be under the Intellectual Property Rights of the researcher. It also excludes ‘Digital’ to avoid restriction to a single storage medium.

Research increasingly involves generation of new NSD and downloaded NSD from public databases. While we have no examples of bilateral benefit-sharing on a contractual basis for use of NSD, the prevailing model of scientific publication of research results and the underlying data means that these results, and NSD, are available globally, the NSD being open access. We have seen evidence that users in all countries are accessing these data. This global availability of information to assist countries in implementing the Convention on Biological Diversity has been called for in a number of COP decisions and under Aichi Target 19.

We are aware that countries do not all have sufficient capacity to make full use of NSD. We regularly engage in capacity building through training and joint research, and see this as a continuing activity. We are concerned that the development of restrictions on use of NSD will damage biodiversity research. Biodiversity loss is alarming⁷⁶, and further restrictions to identify and understand Biodiversity will generate massive drawbacks for the people’s well-being and all life on Earth and consequently endanger research in the conservation and sustainable use of Biodiversity.

Third World Network

SUBMISSION

In response to the Executive Secretary’s invitation pursuant to Decision 14/20, paragraph 9, Third World Network is pleased to submit the following to submit views and information a) to clarify the concept, including relevant terminology and scope, of digital sequence information on genetic resources and if and how domestic measures on access and benefit-sharing consider digital sequence information on genetic resources, and b) on benefit-sharing arrangements from commercial and non-commercial use of digital sequence information on genetic resources.

General

Technological advances in a cluster of linked sciences and technologies including gene synthesis, gene editing, cell culture, epigenetics, genomics-guided breeding, high throughput phenome/genome screening, and numerous other “-omics” applications are, indisputably, changing the ways in which genetic resources are used. Some of these technologies may prove useful for conservation and sustainable use (the first and second objectives), but it is with respect to access and benefit sharing (the third objective) that, for the foreseeable future, the heaviest impacts will be felt.

Without doubt, the technological reality for some years now is that an increasing number of genetic resources can be accessed as digital sequence information (DSI),⁷⁷ rather than as biological material, and

⁷⁶ <https://www.ipbes.net/news/media-release-biodiversity-nature%E2%80%99s-contributions-continue-%C2%A0dangerous-decline-scientists-warn>

⁷⁷ Dormitzer et al. 2013. Synthetic Generation of Influenza Vaccine Viruses for Rapid Response to Pandemics. Science Translational Medicine, 15 May. DOI: 10.1126/scitranslmed.3006368

that users can thereby avoid benefit sharing, both by synthesizing materials from sequences⁷⁸ and by using the DSI itself for commercial purposes.⁷⁹

For example, the pharmaceutical company Regeneron recently used the gene sequence of a Guinean Ebola virus collected in 2014 to create a treatment. Called REGN-EB3, the Ebola drug directly arises from use of Guinean genetic resources. It has received US \$400 million in research support and product orders from the US government. To make the drug, Regeneron downloaded the Guinean sequence information from GenBank and then synthesized portions of the C15 strain genome, using them to generate the monoclonal antibodies of which the patented therapy consists.

The C15 gene sequence was placed in GenBank – which makes sequences available for “free”, no strings attached - by the Nocht Institute, a research center in Hamburg, Germany. While Nocht loaded the C15 sequence into a “free” database, for transfers of samples of Ebola viruses, Nocht uses a legally-binding material transfer agreement (MTA) noting Convention and Nagoya Protocol obligations, and requiring negotiation of a benefit sharing agreement with Guinea in the event of commercial use.

Thus, by downloading the C15 sequence from GenBank and then synthesizing it, rather than requesting a virus sample from Nocht, Regeneron did not sign the MTA requiring negotiation of a benefit sharing agreement, and the company is manufacturing the product for drug stockpiles in North America without benefit sharing to Africa and Guinea.⁸⁰

Some regions and countries anticipated such developments and have access and benefit sharing (ABS) rules on information, and others are now taking policy and legal steps. More generally, however, typical “classic” ABS approaches, often centered on MTAs, are inadequate to ensure benefit sharing for DSI, and ABS approaches are in need of major updates.

The task of creating a system that ensures benefit sharing when genetic resources are utilized as digital sequence information (DSI), and synthesized from DSI, is an existential challenge to the Convention. If benefit sharing is not required of users of DSI, then effective implementation of the third objective of the Convention will prove impossible. Users, especially commercial users, will avoid benefit sharing, thereby undermining the Nagoya Protocol and the Convention itself.

Other international processes related to biodiversity and the Convention look to the CBD for guidance on DSI. These processes, on genetic resources for food and agriculture and pathogens, have important ramifications for biodiversity that is used in public health and agriculture. The progress of these processes may depend on coordination with the Convention’s DSI approach and signals the importance of the Convention and Protocol rapidly moving forward at the next COP/COP-MOP.

Moreover, speed is of the essence because of technology-related concerns. DSI, particularly genetic sequence data, is presently being accumulated at a rate faster than it can be fully leveraged. This is because the ability to sequence is maturing more rapidly than the many ways to use sequence data emerging from genomics and other sciences. As a result, for some commercial purposes, accumulating sequences is akin to depositing money in the bank, or filling a petrol tank – it can be confidently foreseen that even more future uses will emerge for a resource gathered now.

But once such data is sequestered in a private databases, or placed in the (alleged) “public domain” through irresponsible forms of “open access” that do not respect the rights of genetic resource providers, the ability of Parties and IPLCs to protect their sovereign and traditional rights is impaired. The longer it takes for the CBD to act on DSI, the greater the problem becomes.

⁷⁸ Hammond E 2019. Ebola: Company avoids benefit-sharing obligation by using sequences. Third World Network. May. URL: http://www.twon.my/title2/intellectual_property/info.service/2019/ip190504.htm

⁷⁹ See, for example, the “CropOS” product of Benson Hill Biosystems, funded by Google Ventures, which relies on machine-based analysis of a variety of public and proprietary DSI. The system is sold to biotechnology and plant breeding institutions. URL: <https://bensonhillbio.com/design-better-crops-together-benson-hillbiosystems/crop-improvement-platform-cropos/>

⁸⁰ See Hammond 2019, *ibid*, for more information.

Terminology and Scope

With respect to terminology, we note that it is agreed that “Digital Sequence Information” (DSI) is currently used as a placeholder term. Importantly, what the term is agreed to include, that is, its “scope”, is more significant than the words of the term itself. Adapting or even creating a term *de novo* and assigning a definition to it is within the normal scope of implementation of binding agreements, thus, terminology *per se* is not the trouble here. Rather, it is agreement on the breadth of the term.

Whatever term(s) is/are ultimately used, DSI should necessarily include DNA and RNA sequences in all their forms, including assembled and annotated genomes and partial sequences, as well as sequences of alternative forms such as cDNAs, codon optimized sequences, etc. DSI should also include amino acid sequences, SNPs, STR counts, and epigenetic and molecular characterization information (e.g. structures, DNA methylation, etc) and sequence associated metadata (e.g. ‘passport’ data, phenome-genome data, etc).

We note that a useful contribution to the discussion of the various elements of DSI in the context of agriculture - many of which can also be related to biodiversity more generally - can be found in the *Draft Exploratory Fact-Finding Scoping Study on “Digital Sequence Information” on Genetic Resources for Food and Agriculture* prepared for the Commission on Genetic Resources for Food and Agriculture in 2018.⁸¹

We concur with others who believe that the word “digital” is both unnecessary and potentially confusing. It should not be part of the term(s) that is/are ultimately used, due to developments such as quantum computing and the much older phenomenon of storing sequence and other relevant information on non-digital media, such as plain old paper.

We understand DSI as a neologism that originally combined the phrases “digital sequences” with “genetic sequence information”, that were used in the 2015 meeting of the CBD’s Synthetic Biology AHTEG. At that meeting and in other early discussions, use of the word “digital” typically was not focused on the storage format of information, but rather sought to capture the comparative speed and ease with which modern information technology allows sequences to be transmitted across borders, often via the “digital” Internet. DSI is then stored, used *in silico*, and/or subsequently synthesized, and these phenomena pose considerable challenges to methods currently used to implement access and benefit sharing obligations, both in agriculture and beyond.

Thus capturing the above idea – the speed at which large amounts of DSI may be transmitted globally - in the term ultimately used is consistent with how the term arose and would be preferable to the less descriptive “digital”.

Benefit-sharing arrangements from commercial and non-commercial use of DSI

Benefit-sharing arrangements for commercial and non-commercial use of DSI should reflect the same or similar benefit sharing obligations as those attached to biological materials. Benefit sharing obligations should apply to both use of DSI that remains *in silico*, for example, to guide and inform marker-assisted breeding in agriculture, as well as use of DSI to generate materials, for example, vaccine viruses and cell culture strains, or the replication of DSI in material by gene editing.

As is the case with MTAs, *a priori* exemptions from benefit sharing obligations for DSI for categories of users are impractical given the legal structures of non-profit research and intellectual property systems in many countries, particularly developed countries. Rather, benefit sharing for DSI should be structured such that obligations are triggered based on types of use, without regard for the identity of the user, be it a company, a non-profit, an academic, or other entity.

For example, it would be unwise to exempt public or private academic users from benefit sharing obligations when they use DSI because in many countries, as a legal condition of employment, and as a

⁸¹ Heinemann J and D Coray 2018. Draft Exploratory Fact-Finding Scoping Study on “Digital Sequence Information” on Genetic Resources for Food and Agriculture. Commission on Genetic Resources for Food and Agriculture, CGRFA/WG-AqGR-2/18/Inf.10. URL: <http://www.fao.org/fi/staticmedia/MeetingDocuments/AqGenRes/ITWG/2018/Inf10e.pdf>

condition of receiving public funding, academics are required to patent, and their institutions legally obligated to financially exploit, any commercially-applicable discovery, even if that discovery is incidental. Moreover, others may derive commercial uses from DSI that is initially handled by academic entities.

The US Bayh-Dole Act is an example of a law creating the above type of requirements, and Bayh-Dole Act types of legal obligations and institutional policies permeate the developed world (and some developing countries). Where such laws are in place, put simply, it is against policy or even illegal for academics not to patent and commercially exploit academic inventions, even if they are made accidentally.

While aspects of policy and law vary from country to country, across the North, it can consistently be observed that academic users of DSI:

1. Must report all inventions to their employer, even if unintended;
2. Must convey rights to inventions (or proceeds from them) to the employer, either as a matter of law or terms of employment;
3. Receive payments as personal income from their employer in return for their compliance;
4. May not waive or negotiate the intellectual property interests and policies of their institution in an access and benefit sharing agreement.

Having said the above, while all users of DSI should be subject to benefit sharing requirements, not every use of DSI – indeed not the vast majority of uses outside of corporate confines – would necessarily trigger action in relation to those obligations, particularly if the use proves to be genuinely non-commercial. Thus, the benefit sharing obligations incumbent on users of DSI can and should contain carefully crafted thresholds that ensure that genuinely non-commercial research will typically proceed without triggering financial benefit sharing obligations.

One way to cause such benefit sharing obligations to take effect is through the use of **data access and use agreements** that DSI users must agree to before accessing data, for example, DSI in internet-linked databases. Data access and use agreements offer the potential to permit DSI to remain publicly-accessible while protecting the interests of providing countries and IPLCs. Akin to the “terms and conditions” that accompany an airplane or train ticket, or utilities such as water and sewer services, data access and use agreements can be used to set forth benefit sharing obligations connected to publicly-accessible DSI.

Opponents of benefit sharing for use of DSI sometimes suggest, while seldom offering evidence, that data access and use agreements and similar approaches would inherently impede research. Review of online databases related to biodiversity, however, confirms that it is quite normal for databases to employ terms and conditions, and this includes databases specializing in widely diverging sets of biodiversity.

The Global Biodiversity Information Facility (GBIF)⁸² emphasizes “free” data sharing in a way that may be inappropriate for some DSI, however even this “free” database aggregator utilizes a “data use agreement” that “requires and implies agreeing” to a variety of stipulations. These include that the national law of Denmark (GBIF Secretariat host) governs the agreement, recognition of the “scope and application of Intellectual Property Rights and benefit sharing agreements as determined under relevant laws, regulations and international agreements,” that data providers may restrict access to data, and that “reasonable assurances” be made by data providers that prior informed consent has been obtained.

(The latter, and some other stipulations of the GBIF data use agreement, fall well short of protecting the benefit sharing interests of provider countries and IPLCs in DSI, however, the fact that such stipulations exist in the GBIF user agreement demonstrate that terms and conditions relevant to DSI are already in use.)

The Center for Australian National Biodiversity Research and the Australian National Botanic Gardens assert a Commonwealth copyright and other restrictions on data hosted at their websites,⁸³ which include a number of biodiversity databases. These restrictions expressly prohibit commercial use, stating that, unless otherwise noted, “*Information and data on this server have **NOT** been placed in the public domain, but are provided for the personal non-commercial use of educators, students, scholars and the public,*” and

⁸² <https://www.gbif.org/en/terms/data-user>

⁸³ <http://www.anbg.gov.au/copyright.html>

that “*Any commercial use or publication of these resources without a licence from the custodians is strictly prohibited*” (emphasis in the original).

The “World Database of Key Biodiversity Areas”⁸⁴ (“KBA”), managed by Birdlife International on behalf of a consortium including IUCN, WWF, and a number of other organizations employs a “terms and conditions of use” document that asserts copyright and ownership over the data, prohibits commercial use, prohibits reposting or redistribution, and prohibits commercial use of derivative works. This latter restriction has substantial implications with respect to intellectual property.

The Seabird Tracking Database,⁸⁵ also managed by Birdlife, has terms of use that assert copyright and state that database users may not “*adapt, alter or create a derivative work from any BirdLife International content except for your own personal, non-commercial use.*” This and other terms “*shall take effect immediately on your first use*” of the website.

The Marine Metagenomics Portal,⁸⁶ based in Norway, takes a less restrictive approach than Birdlife, IUCN, WWF, and colleagues, utilizing a Creative Commons No Derivatives license. This allows copying and redistribution of the data, however, users may not redistribute the data if they “*remix, transform, or build upon*” it, and users “*may not apply legal terms or technological measures that legally restrict others.*” These restrictions have an effect of limiting any intellectual property claims. To enhance clarity, the terms of use add that “*Any genetic information is provided for research, educational and informational purposes only.*”

The Arabidopsis Information Resource (TAIR)⁸⁷ maintains a database of genetic and molecular biology data for the model higher plant *Arabidopsis thaliana*. This database has lengthy terms that include “*You may not utilize the Service if you are an employee (including part-time employees) or contractor of, or in any way acting on behalf of a for-profit entity.*” Access to portions of the data requires registration and a paid subscription. The terms may be amended at any time without notice to users.

In the public health field, a number of DSI databases linked to pathogens impose terms and conditions on users. In the field of public health, in contrast to claims that benefit sharing for DSI will impede data sharing, in several instances databases with terms and conditions designed to protect the interests of genetic sequence data providers have become lauded examples of sharing.

For instance, GISAID, the Global Initiative on Sharing All Influenza Data, employs a detailed user agreement.⁸⁸ Among its provisions, it prohibits users from reposting data in other databases, states that data providers have not relinquished data ownership by making it available to GISAID users, and prohibits users from providing data to third parties. While GISAID predates the Nagoya Protocol and its user agreement does not require benefit sharing, the highly elaborated GISAID agreement, with its emphasis on protecting data provider interests, has proven very popular among influenza researchers.

MeaNS, the World Health Organization’s database of measles infection data and measles strain gene sequences, operated by Public Health England,⁸⁹ utilizes terms and conditions that prohibit the downloading of sequence data: “...as part of the terms and conditions of site usage, downloading sequences is not allowed. This is to ensure that users who submit sequences will not have sequences used by other individuals without permission.” Some DSI in MeaNS is available in other, less restrictive, databases, but other data are not. The MeaNS terms further require database users to be academic or non-profit researchers and state that users may not use the database for “reproduction, adaptations, preparation of derivative works, or distribution of copies of any portion of MeaNS for any purpose.”

⁸⁴ <http://www.keybiodiversityareas.org/info/dataterms>

⁸⁵ <http://seabirdtracking.org/termsfuse>

⁸⁶ <https://mmp.sfb.uit.no/terms/>

⁸⁷ https://www.arabidopsis.org/doc/about/tair_terms_of_use/417

⁸⁸ <https://www.gisaid.org/registration/terms-of-use/>

⁸⁹ MeaNS Terms and Conditions: http://www.who-measles.org/Public/Web_Front/terms_conditions.php Also see the FAQ: http://www.who-measles.org/Public/Web_Front/faq.php

RubeNS, the World Health Organization's rubella database, also operated by Public Health England,⁹⁰ utilizes similar restrictions.

Thus, for databases hosting DSI of a wide array of biodiversity, including the databases noted here, agreement to terms and conditions is frequently required before access to information is granted, and those terms and conditions frequently impose particulars on commercial use and govern DSI including gene sequences.

It is worthwhile to explore how making access to DSI contingent on acceptance of data access and use agreements, and how such agreements can protect the rights of providers, can be means by which to operationalize benefit sharing for DSI and thereby protect the Convention and advance implementation of its third objective and the Nagoya Protocol.

UK BioIndustry Association (BIA)

SUBMISSION

1. About the BIA

The BIA is the trade association for innovative life sciences in the UK. Our goal is to secure the UK's position as a global hub and as the best location for innovative research and commercialisation, enabling our world-leading research base to deliver healthcare solutions that can truly make a difference to people's lives.

Our members include:

- Start-ups, biotechnology and innovative life science companies
- Pharmaceutical and technological companies
- Universities, research centres, tech transfer offices, incubators and accelerators
- A wide range of life science service providers: investors, lawyers, IP consultants, IR agencies

We promote an ecosystem that enables innovative life science companies to start and grow successfully and sustainably.

For any further information on the contents of this submission please contact Eric Johnsson, Policy and Public Affairs Manager, by emailing ejohnsson@bioindustry.org.

2. Executive Summary

The BIA and its members do not believe digital sequence information ("**DSI**") should be included within the scope of the objectives of the Convention on Biological Diversity (the "**CBD**") and the objective of the Nagoya Protocol ("**NP**", as further defined below) on both legal and practical grounds.

The BIA and its members are concerned about the focus of this consultation given that compliance with the NP is still in its infancy. Indeed, addressing ongoing compliance and implementation challenges of the NP should be the main focus of the Parties to the NP and of the Secretariat to the CBD. The potential inclusion of DSI at this stage would complicate matters further, exacerbating the significant and complex issues and challenges.

Moreover, inclusion of DSI would do more harm than good by, amongst other things:

- presenting additional compliance challenges and problems which could seriously stifle innovation, particularly for SMEs; and
- resulting in unintended consequences on the country of origin of the underlying genetic resource ("**GR**").

⁹⁰ <http://www.who-rubella.org/tanc>

As DSI is not a GR, any measures to include DSI within the scope of the objectives of the CBD and NP would require the NP to be formally amended.

Reaching a satisfactory definition for DSI would be very challenging. The BIA queries why significant resources are being incurred in determining the scope of the definition of DSI and the mechanism for access and benefit-sharing ("ABS") arrangements when DSI does not fall within the scope of the CBD or NP and consensus has not been reached as to its inclusion in principle.

Given that measures under the NP are meant to be "*appropriate, effective and proportionate*", careful consideration needs to be given as the due diligence challenges to the inclusion of DSI within the scope of the NP would be disproportionately burdensome.

3. Background

The Convention on Biological Diversity (the "**CBD**") entered into force on 29 December 1993 and has three main objectives:

1. The conservation of biological diversity;
2. The sustainable use of the components of biological diversity; and
3. The fair and equitable sharing of the benefits arising out of the utilization of genetic resources ("**GRs**") (including by appropriate access to GRs and by appropriate transfer of relevant technologies, taking into account all rights over those resources and to technologies, and by appropriate funding).

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 (the "**NP**") is a supplementary agreement to the CBD. It intends to provide a transparent legal framework for the effective implementation of the third objective of the CBD: the fair and equitable sharing of benefits arising out of the utilization of GRs. It was adopted at the tenth meeting of the Conference of the Parties, in Nagoya, Japan on 29 October 2010 and entered into force on 12 October 2014.

In the EU, the NP has been implemented under [Regulation \(EU\) No 511/2014](#) of 16 April 2014 (the "**EU NP Regulation**") and accompanying Commission [Implementing Regulation \(EU\) 2015/1866](#) of 13 October 2015.

In December 2016, the Conference of the Parties to the CBD adopted a decision to consider any potential implications of the use of digital sequence information ("**DSI**") on GRs on the three objectives of the CBD ([decision XIII/16](#)). The Conference of the Parties serving as the meeting of the Parties to the NP invited submissions to include information relevant to the NP ([decision NP-2/14](#)). The Secretariat to the CBD subsequently invited the submission of views and information through [notification 2017-37](#). The BIA submitted a response to this consultation in September 2017⁹¹ (the "**BIA 2017 Response**"), as did many other stakeholders⁹². Due to the interrelatedness between this response and the BIA 2017 Response, we have attached a copy of our previous response and recommend that both be read in tandem.

In November 2018, the Conference of the Parties to the CBD adopted [decision 14/20](#) to establish a science and policy based process on DSI. On 5 February 2019, pursuant to decision 14/20 paragraph 9, the secretariat to the CBD in a [notification](#) invited parties to the CBD, other Governments, indigenous peoples and local communities, relevant organizations and stakeholders to submit views and information:

- (a) To clarify the concept, including relevant terminology and scope, of DSI on GRs and if and how domestic measures on access and benefit-sharing ("ABS") consider DSI on GRs;
- (b) On benefit-sharing arrangements from commercial and non-commercial use of DSI on GRs.

Pursuant to decision 14/20, paragraph 10, the secretariat to the CBD also invited parties to the CBD, other Governments and indigenous peoples and local communities to submit information on their capacity-

⁹¹ <https://www.bioindustry.org/resource-listing/bia-response-dsi-regulation-in--nagoya.html>

⁹² <https://www.cbd.int/abs/dsi-gr/ahteg.shtml>

building needs regarding the access, use, generation and analysis of DSI on GRs, in particular for the three objectives of the CBD.

4. Incorporation of DSI within the scope of the CBD and Nagoya Protocol

Whilst the BIA and its members support the three objectives of the CBD and the objective of the NP, we do not believe DSI should be included within the scope of the objectives of the CBD and the objective of the NP on both legal and practical grounds. Indeed, when our members were consulted in 2017 on how they might be impacted by the proposed incorporation of DSI into the NP **all responding members strongly disagreed with the proposed incorporation of DSI into the NP.**

4.1. Compliance with the NP is still in its infancy

On 24 January 2019, the European Commission published the first [report](#) on the implementation of the EU NP Regulation. The report is mainly based on information from national reports submitted by all 28 EU Member States. It includes the following conclusions as to the status of play and the identified challenges:

- The implementation of the EU NP Regulation is still in its early days with many Member States starting relatively late to set up necessary institutional and administrative frameworks, and with implementation and enforcement being slow and uneven amongst Member States;
- Lack or limited human and financial resources devoted to the implementation of the EU NP Regulation is often reported as a major obstacle;
- Lack of specialized personnel and qualified experts is also identified as a problem;
- Concerns have been raised by Member States as to the significant administrative burden and costs of implementing the EU NP Regulation;
- Delays in designating Competent Authorities has slowed down the implementation of other provisions of the EU NP Regulation;
- A low level of interest, among Member States, in becoming a registered collection in the EU register of collections;
- A low level of awareness among Member States institutions and administrations, as well as by stakeholders about the obligations stemming from the NP and EU NP Regulation; and
- Interpretation challenges with requests for further guidance to clarify some of the terms and more real examples on implementation to clarify the issues.

According to the UK report on the implementation of the EU NP Regulation prepared by the Department for the Environment, Food and Rural Affairs ("DEFRA"), as at 31 August 2017, no due diligence declarations had been received based on Articles 7(1) or 7(2) of the EU NP Regulation. The lack of such declarations is reported to be due to the low level of awareness among stakeholders and, in many cases, GRs being accessed before 12 October 2014 (when the EU NP Regulation came into force) still being used in R&D. DEFRA had yet to conduct checks on users at the time the report was published with the first check stated in the report to be planned in October 2017.

Given that compliance with and implementation of the NP is still in its infancy with numerous challenges still needing to be addressed, the BIA and its members are concerned about the focus of this consultation. Indeed, the BIA and its members consider that addressing ongoing compliance and implementation challenges of the NP as it currently stands should be the main focus of the Parties to the NP and of the Secretariat to the CBD. The potential inclusion of DSI at this stage would complicate matters further, exacerbating the significant, complex issues and challenges that have already been identified.

Further, as we stated in the BIA 2017 Response, *"any decision to amend the scope of the NP must be based on clear evidence that the ABS objective is not working, the current ABS system is failing and that the incorporation of DSI would help achieve the ABS objective and remedy the identified failing in the ABS system. We are not aware of such evidence."*

If there are shortcomings with the ABS system, it is largely due to the lack of provider country laws which facilitate access and thus generate benefits. Once addressed, and comprehensive legal frameworks of

national ABS laws are put in place, concerns about the lack of benefit sharing related to genetic resources access and use should disappear."

4.2. Inclusion of DSI within the NP would do more harm than good

The BIA and its members believe that the inclusion of DSI would do more harm than good by presenting additional compliance challenges and problems which could seriously stifle innovation. The BIA already has evidence from its members that the NP is having a negative impact on R&D. Due to uncertainties as to the exact nature and scope of the obligations to be fulfilled, many stakeholders are putting measures in place to navigate the NP in such a way as to mitigate against any disruption to innovation as further evidenced below:

AstraZeneca

In a [presentation](#) made by AstraZeneca on 22 February 2018 at the BIA Committee Summit, it described its approach to proactively engaging with compliance under the NP by, amongst other things:

- Establishing a Nagoya Governance Team
- Defining the Company's public policy position
- Modifying its Bioethics Policy
- Developing a Global Standard defining individual responsibilities
- Developing a Nagoya Sourcing e-tool to determine if GRs are in or out of scope of the NP
- Creating 3-minute training videos that provide an overview of the NP and the Company's responsibilities

AstraZeneca has the resources to implement sophisticated measures in order to comply with the NP but not all biotech companies are as well resourced.

Prokarium

Prokarium is developing a technology platform based on engineered bacteria to prevent infectious diseases and treat solid tumours. It ensures compliance with the NP by selecting microorganisms from a non-NP country. As a result, the regulation is having unintended consequences of reducing ABS with signatory source countries due to difficulties associated with compliance; inclusion of DSI will only exacerbate this.

We also refer to the BIA 2017 Response which considered the following three key points in some detail and which exemplified some of the negative consequences of incorporating DSI into the NP for our members:

1. The incorporation of DSI into the NP will lead to further legal uncertainty and compliance difficulties for SMEs;
2. DSI regulation in the NP will hinder SMEs' R&D; and
3. The incorporation of DSI into the NP poses serious public health concerns.

Moreover, the inclusion of DSI into the NP could have unintended consequences on the country of origin of the underlying genetic resource. Open access to and use of DSI is widely accepted to be a fundamental driver of scientific research and innovation. Limiting or hindering access to DSI by requiring ABS agreements to be entered into before being entitled to access the information would in all likelihood have a negative impact on the conservation and sustainable use of components of biological diversity (two objectives of the CBD). This is because such DSI would be less likely to be accessed and used for research and development. This would undermine the potential to conserve and sustainably use the underlying genetic resources which would, in turn, negatively impact the country in which such genetic resource is found.

In the next section we set out some additional points particularly pertinent to the current consultation and re-state previous points which we consider important to raise in the context of this consultation.

5. DSI is not a genetic resource

The current consultation appears to be based on the erroneous assumption that DSI falls within the scope of the CBD and the NP.

According to the CBD, GRs means "genetic material of actual or potential value" and genetic material means "any material of plant, animal, microbial or other origin containing functional units of heredity". DSI cannot be genetic material as it is merely a representation of the sequence of a biological molecule (e.g. DNA). Since it is not a physical material (whether plant, animal, microbial or other form that contains functional units of heredity), it cannot be a GR as such.

The NP incorporates the CBD definition for GRs and genetic material and also defines a derivative as "a naturally occurring biochemical compound resulting from the genetic expression or metabolism of biological or genetic resources, even if it does not contain functional units of heredity". DSI is information and, as such, cannot be said to be "naturally occurring" or a "biochemical compound" let alone that it (the information) can result from the "genetic expression or metabolism of biological or genetic resources".

Consequently, DSI cannot be legitimately brought within the scope of the CBD and the NP without first amending the CBD and/or the NP itself. Any such amendment would require agreement by consensus from the parties to the CBD and the NP or, as a last resort, be adopted by a two-thirds majority vote of the parties to the instrument in question.

6. What is DSI?

Asking stakeholders to define DSI for the purposes of ABS arrangements requires stakeholders to consider issues in the wrong order. If information (which is what DSI is) cannot fall within the scope of the CBD or NP because it is not a genetic material, GR or derivative and consensus has not been reached with parties to the CBD and the NP as to its inclusion in principle, the BIA queries why significant resources are being incurred in determining the scope of the definition of DSI and the mechanism for ABS arrangements. This is not least because the implementation of the NP as it currently stands it still in its infancy (see section 4.1 above).

Moreover, attempting to define DSI is not straightforward and raises more questions than answers. Indeed, the term "digital sequence information" is broad in scope and does not encompass a single type of data:

- If you limit DSI to DNA, it would not include RNA genomes (such as a retrovirus). Should DSI be limited to genomic DNA or RNA sequences?
- Should it be limited to native DNA (i.e. the form found in nature) or should it include only the coding regions?
- What about regulatory DNA that does not code for proteins but has other effects (e.g. processing genes)?
- DSI is often edited, codon-optimised or compiled from alignments of other sequences. What happens when you modify DNA sequences *in silico* or create a compilation of synthetic DNA from different sources to develop novel molecules and functions – would they still fall within the scope of the definition for DSI?
- If an 'NP sequence' forms part of an alignment, is the resulting consensus sequence subject to the obligations under the NP?
- As DNA/RNA sequencing technology is not 100% error-free and there is a high degree of natural variation in genetic sequences within populations, what level of alignment would be required to invoke obligations under NP?
- What if only a partial sequence is known?
- What about information relating to the secondary or tertiary structure of DNA?
- What about any annotations to DSI?

7. Measures on access and benefit sharing for DSI

Given that DSI does not fall within the scope of the CBD or the NP, in the UK there are no measures and arrangements on ABS for DSI on GRs whether commercial or non-commercial.

8. Due diligence challenges

As we reported in the BIA 2017 Response, DSI has been generated, stored and used for several decades in vast and increasing quantities. As at 2015 it was estimated that "*publicly available databanks now contain quadrillions ($>10^{15}$) of nucleotides of DNA sequence data, soon to be quintillions ($>10^{18}$ bases). These have been collected from over 300,000 different species of organisms*"⁹³.

The GenBank sequence database is one of many such DNA sequence databanks. It is an open access, annotated collection of all publicly available nucleotide sequences and their protein translations which is produced and maintained by the US National Center for Biotechnology Information as part of the International Nucleotide Sequence Database Collaboration. From 1982 to the present, the number of bases in GenBank has doubled approximately every 18 months⁹⁴. As at April 2019, the number of base pairs recorded on GenBank were 321,680,566,570.

The country of origin of DSI is not always recorded and traceability of sequences would be a key challenge if DSI were to be included within the scope of the NP. Natural variation and mutations that occur over time also clouds traceability. Moreover, as an organism can often exist in multiple countries, questions arise as to which country should derive benefit from the DSI of that organism (especially if the original country of origin for that particular sequence information is unknown).

This, together with the sheer size of data being generated and stored, would create additional due diligence challenges and would be disproportionately burdensome should DSI eventually fall within the scope of the NP. Given that measures under the NP are meant to be "*appropriate, effective and proportionate*" careful consideration needs to be given to the basis for the justification of extending the scope of the NP beyond GRs and into DSI.

<p>Wellcome Sanger Institute</p>

SUBMISSION

Introduction

1. The Wellcome Sanger Institute uses genomic sequences to advance the understanding of human and pathogen biology to improve human health. We use science at scale to tackle the most challenging global health research questions.
2. The Wellcome Sanger Institute, based on the Wellcome Genome Campus, Hinxton, UK, is a world-leading hub for genomes and biodata research. The campus is also home to EMBL-EBI, Connecting Science, Sanger Institute spin-out companies, start-up companies and Genomics England.

Key messages

3. We strongly disagree with the proposal to include digital sequence information (DSI) in the scope of the Convention on Biological Diversity (CBD) and the Nagoya Protocol.
4. We fully agree that countries should share equitably in the benefits of research and development which utilises sovereign genetic resources, but it is our view that the inclusion of DSI in the Nagoya Protocol would fail to achieve this goal and has the potential to do more harm than good.
5. Science is a global endeavour and the ability to tackle complex global challenges depends on international collaboration and the ability to share and freely access research findings and data. International research, development and surveillance activities could be seriously threatened by the inclusion of DSI in the Nagoya Protocol.

⁹³ Pevsner, J. (2015). *Bioinformatics and functional genomics*. Chichester, West Sussex: Wiley Blackwell.

⁹⁴ <https://www.ncbi.nlm.nih.gov/genbank/statistics/>

6. Globally, researchers are sharing and utilising DSI on animals, plants and pathogens via open access databases. It is absolutely vital that free and unrestricted sharing of DSI continues. To impose terms and conditions or access agreements would seriously undermine the value, accessibility and sustainability of DSI.

The value of DSI

7. DSI holds tremendous value for understanding biology and evolution. The development of genomics is a direct result of DSI being freely available and shared globally in easy-to-access public databases. This widespread sharing of DSI is fundamental for advancing research and driving innovation, for example, by addressing conservation challenges, delivering actionable public health strategies and responding rapidly to global public health emergencies.
8. Public health emergencies require rapid responses. The rapid and unconstrained availability of DSI from genetic resources is critical for quickly determining how the disease is spreading, how the pathogen is evolving and identifying the genes involved in disease onset. This information is vital for informing emergency response tactics and public health strategies. This process is time-sensitive and delays due to the establishment and negotiation of licencing agreements on a bilateral basis could be disastrous for public health.

Global accessibility of DSI in easy-to-access databases

9. The Sanger Institute shares non-human DSI openly via the European Nucleotide Archive (ENA), which is maintained by EMBL-EBI. As of 2018, the ENA comprised a comprehensive database of 1.5 billion sequences and this number continues to grow⁹⁵. Databases like the ENA collate, integrate, curate and make freely available scientific data from around the world. DSI within these databases is easily findable (via accession numbers) and accessible to researchers and publics around the world. The ENA offers a variety of support programmes and training packages to facilitate the use of DSI. This allows researchers who do not have access to the infrastructure and resources required for large-scale sequencing projects to still benefit from the resource.
10. Open access databases are founded upon the principle of open science. They typically have an intricate and integrated network of services, tools and data resources that exist to support the discovery, analysis and interpretation of DSI. Genome sequences are compared to thousands of other genomic sequences from different organisms, using tools such as BLAST, Pfam and InterPro, to identify features such as genetic regions, conserved genomic domains and protein structures. These comparisons of thousands of sequences provide insights into characteristics including virulence, metabolism and genetic lineage. The true value of DSI comes from aggregated datasets where patterns of conserved features or differences can be easily identified. The unconstrained aggregation of thousands of genome sequences is fundamental to providing biological insights in genetic resources.
- 11. Restrictions on the open sharing of DSI in databases like the ENA and its affiliated network could create a major barrier to research and innovation across a whole array of research fields.**

Open access science

12. Since the 1990's, there has been a push towards open access science and genomics has been a global leader⁹⁶. Initiatives like Plan S (a framework for transitioning to fully open access publishing)⁹⁷ are bringing funders together to ensure that research findings and associated data are freely and openly available to other researchers and the general public.

⁹⁵ Harrison, P.W. et al. (2019). The European Nucleotide Archive in 2018. *Nucleic Acids Research*, 47, D84-D88. <http://europepmc.org/abstract/MED/30395270>

⁹⁶ Sharing Data from Large-scale Biological Research Projects: A System of Tripartite Responsibility, Report of a meeting organised by the Wellcome Trust and held on 14–15 January 2003 at Fort Lauderdale, USA.

⁹⁷ <https://www.coalition-s.org/>

13. Open access to research data brings a far greater return on investment for taxpayers, provides researchers with the opportunity to work with data they may not have the capacity to generate themselves and circumvents the unnecessary duplication and substantial cost of reproducing DSI.
- 14. Restrictions imposed on sharing datasets openly, by incorporating DSI into the Nagoya Protocol, will reduce accessibility to science, increase the cost of science and reduce the return on investment to the public.**

Restricted data

15. Researchers are more likely to use resources, such as DSI, when access is not constrained by payment or access agreements.
16. The Sanger Institute shares data free of charge, but access to the majority of our human DSI requires a data access agreement because of the sensitivities around human genomic data. Anecdotally, we know that the requirement to sign a data access agreement can be a barrier to access, particularly for researchers in some low and middle income countries.
17. Development of data access processes is non-trivial, requiring dedicated sustainable resource to maintain them and more sophisticated researcher environments to use them.
- 18. Imposing additional regulation and/or requirements for access agreements will discourage and limit use of that genetic resource, risking failure to achieve the goals and ambitions of the CBD.**

Access and benefit sharing arrangements

19. Existing Nagoya Protocol mechanisms are already impacting research at the Sanger Institute. Established research projects with long-term collaborators in Ethiopia and Gabon have a track record of training and capacity building, but have stalled due to uncertainties around the existing Nagoya Protocol. The inclusion of DSI within scope of the Nagoya Protocol will exacerbate these issues, cause delays to other research projects and hinder scientific progress.
20. International collaborations are imperative to our ability to tackle complex global challenges. We frequently collaborate with researchers in host countries when generating DSI from their sovereign genetic resources. These successful collaborations are beneficial to both parties for advancing research and scientific understanding as well as boosting recognition and capacity building efforts in host countries.
21. Countries exercising sovereign rights over their genetic resources are able to benefit from access and benefit sharing (ABS) arrangements in accordance with the existing Nagoya Protocol. Given that the generation of DSI requires physical samples of genetic resources, ABS can be achieved using existing Nagoya mechanisms.
22. The inclusion of DSI within the scope of the Nagoya Protocol will impose significant barriers to research and innovation and could ultimately risk inequitable ABS. For example, if DSI from a genetic resource is released freely and openly by one country, but the DSI from a very closely related genetic resource is constrained by its respective country, then investment and research is more likely to focus on the unconstrained, rather than constrained, genetic resource. In turn, resulting research outputs might only be relevant to specific geographical regions where DSI is freely available.
- 23. Equitable access and benefit sharing can be achieved under the existing framework of the Nagoya protocol.**

Case Study 1 - Darwin Tree of Life Project

Reference genomes provide a template for assembling and comparing individual genomes of the same species. As of October 2017, the National Centre for Biotechnology Information (NCBI) database held 2,534 sequenced eukaryotic species genomes, which represents less than 0.2% of all known eukaryotes⁹⁸.

The Darwin Tree of Life project, launched in November 2018, seeks to sequence and create reference genomes for the 66,000 known eukaryotic species in the British Isles over the next 10 years. Led by the Wellcome Sanger Institute, this extensive project will draw on the expertise of several UK organisations, including the Natural History Museum, Royal Botanic Gardens Edinburgh, Earlham Institute, Universities of Cambridge, Oxford and Edinburgh, EMBL-EBI, Marine Biology Association and others for sample collection, DNA sequencing, genome assembly and annotation, and data storage. The project will feed in to the wider Earth BioGenome Project - a global collaborative effort to sequence the genomes of all 1.5 million known species of animals, plants, protozoa and fungi on Earth.

The Darwin Tree of Life and Earth BioGenome projects will revolutionise our understanding of biology and evolution. However, building such an urgently needed reference dataset for the planet's dwindling species list will require a multi-billion dollar global effort with massive automation of sample processing and sequencing. The subsequent sequence data needs to be stored in public domain databases and made freely available for wider research use. The project can only achieve its scientific goals of protecting biodiversity, supporting conservation and understanding the ecosystems around us if the data are used globally and by future generations.

The inclusion of DSI within scope of the Nagoya Protocol risks hindering the development of comparable projects in other countries and thus impeding access to the benefits of such endeavours. Any attempt to prohibit rapid and widespread sharing of invaluable DSI would prevent researchers and conservationists globally from advancing their research and responding promptly to existing and emerging threats to public health and biodiversity.

Case study 2 – Tracking the spread of cholera at the household level – Prof Nicholas Thomson

Every year, cholera causes up to 143,000 deaths and the disease-causing bacteria, *Vibrio cholerae*, infects up to 4 million people worldwide. Cholera is caused through the consumption of food or water contaminated with *Vibrio cholerae* and the disease remains to be a global public health threat. The capital city of Bangladesh, Dhaka, is a megacity and is hyper-endemic for cholera; experiencing two seasonal outbreaks every year. Very little was previously known about the diversity of the *Vibrio cholerae* strains that circulated around Dhaka and their role in the seasonal outbreaks. For the first time, to understand the diversity and transmission of cholera in an endemic setting, Wellcome Sanger Institute researchers and their collaborators tracked the transmission of cholera at the household level.

Vibrio cholerae was isolated from index patients who had been admitted to the International Centre for Diarrhoeal Disease Research, Bangladesh (icDDR,b) and over a period of 3 weeks, follow-up samples were collected from household contacts who shared a cooking pot with the index patients.

By sequencing these 303 *Vibrio cholerae* isolates from 224 individuals across 103 households, the researchers were able to determine how the cholera strains from each person were related and then compare them with other strains from around the world.

They found that nearly 80% of secondary infections were linked to the first index case within the same household, which implied that once cholera had entered the household, it spread between household members rather than repeatedly coming in from outside.

These findings strongly reinforce the importance of sanitation and hygiene to prevent the chain of transmission within households. These findings can be used by public health officials to improve cholera

⁹⁸ Lewin, H. A. et al. (2018). Earth BioGenome Project: Sequencing life for the future of life. PNAS, 115, 4325-4333. <https://www.pnas.org/content/115/17/4325>

control strategies, for example, improving sanitation, water chlorination and vaccinating household members to help reduce the spread of this deadly disease.

The inclusion of DSI in the scope of the Nagoya Protocol could cause prohibitive delays to research projects like this and hinder rapid public health efforts to control disease epidemics.

Case study 3 – Mutational signatures as a tool for health research and conservation - Dr Alex Cagan

Mutational signatures as a tool for health research

Nearly 990 people in the UK are diagnosed with cancer every day and 1 in 2 people will be diagnosed with a form of cancer during their lifetime⁹⁹. Cancer is caused through the accumulation of mutations in the DNA of somatic tissue. It might be expected that larger animals or those with longer lifespans should have a higher incidence of cancer, but this trend is not seen in nature.

In an ongoing project at the Wellcome Sanger Institute, researchers are comparing somatic mutation rates across a variety of different species. In collaboration with the Zoological Society of London (ZSL), UK Cetacean Strandings Investigation Programme (CSIP), the University of Cambridge and others, researchers are using DSI, histology and laser-capture microscopy to estimate mutation rates and identify specific mutational signatures across different species. This research will provide insights into the evolution of mutation rates across different species and help us to further understand cancer and aging in humans.

Researchers are particularly interested in animals with long lifespans and low mutation rates as well as those that have mechanisms for reducing their mutation rate. Studying these animals will help us to understand the somatic mutation rate in humans and could identify ways of reducing these rates to prevent human cancers. Many of these animals, however, are not native to the UK or EU and, as such, researchers have currently delayed their focus on these animals as a result of the uncertainties around the requirements of the existing Nagoya Protocol. Inclusion of DSI risks researchers excluding species from certain countries and exacerbating existing inequalities.

Mutational signatures as a tool for conservation

Polluting and toxic chemicals in the natural environment can cause an array of health concerns in both humans and in animals. Characteristic mutational signatures can provide insights into the threats posed to endangered animals.

Polychlorinated biphenyl (PCB) is a highly toxic industrial compound known to cause cancer and developmental damage. Although its use is now widely prohibited, it is still persistent in the environment through spills, leaks and improper disposal and storage of industrial equipment. Like other carcinogens, PCB causes unique mutational signatures that enable researchers to identify specific environmental threats endangering different animals.

Biomonitoring of habitats using mutational signatures will help determine the impact on health and the threats posed to species by environmental change. This will help inform specific environmental management and conservational measures to prevent further endangerment of animals. This type of research will only be effective if DSI can be openly shared.

Case study 4 – Pathogen surveillance – Professor David Aanensen

Antimicrobial resistance is an increasingly serious threat to global public health that requires international action. Pathogenic bacteria infect millions of people worldwide every year and are responsible for many diseases including tuberculosis, pneumonia and typhoid. Many pathogenic bacteria have developed drug resistance and subsequently several antibiotics are no longer effective against these life-threatening diseases. Pathogens do not respect geographical borders and antimicrobial resistance is present in every country.

⁹⁹ <https://www.cancerresearchuk.org/health-professional/cancer-statistics-for-the-uk>

The Centre for Genomic Pathogen Surveillance (CGPS), based at the Wellcome Sanger Institute and led by Professor David Aanensen, performs global surveillance of pathogens using whole genome sequencing to understand the emergence and spread of diseases and drug resistance.

The CGPS combines structured population surveys and whole genome sequencing to identify specific pathogenic strains present in different regions and how they can spread within and between countries. This form of genomic technology is becoming the gold-standard tool for surveillance and is used to inform public health policies.

The surveillance data generated by the CGPS is freely and openly available in public repositories (e.g. the ENA) and the team is continually developing easy-to-use software tools to make data integration, visualisation and interpretation accessible to all. The CGPS invests in capacity-building in low- and middle-income countries to help train the future leaders of new national and emerging surveillance programmes.

Rapid sharing of surveillance DSI is vital for tracking and anticipating the geographical routes of diseases and drug resistance and for informing regional and national patient treatment strategies to prevent drug resistance spreading to other regions.

Pathogen surveillance requires research to be performed across geographical borders.

Case study 5 – Conservation of endangered gorilla populations – Dr Chris Tyler-Smith

Mount Tshiaberimu in the Democratic Republic of Congo (DRC) is home to a highly endangered population of gorillas. With a population of one female, four males and one baby of unknown sex, the colony is no longer viable and requires new gorillas to provide enough genetic variability for the colony to continue.

Living on Mount Tshiaberimu, they were thought to be mountain gorillas, but some unusual characteristics put this in to question. To confirm their species, researchers at the Wellcome Sanger Institute alongside their US collaborators and conservationists from the DRC took a sample from a gorilla, which was sequenced at the Sanger Institute and then compared to openly available DSI from three gorilla species. The analysis showed that it was, in fact, not a mountain gorilla, but an eastern lowland subspecies. This information was necessary for developing a conservation strategy.

Identifying the gorilla species was only possible because the researchers were able to access DSI of the other gorillas.
