Received: 23 December 2020 | Revised: 24 March 2021 | Accepted: 26 March 2021

DOI: 10.1002/ppp3.10201

Plants People Planet PPP Open Access

OPINION

Practical consequences of digital sequence information (DSI) definitions and access and benefit-sharing scenarios from a plant genebank's perspective

Martin Brink | **Theo van Hintum**

Centre for Genetic Resources, the Netherlands (CGN), Wageningen University & Research, Wageningen, the Netherlands

Correspondence
Martin Brink, Centre for Genetic Resources, the Netherlands (CGN), Wageningen University & Research, Wageningen, the Netherlands.
Email: martin.brink@wur.nl

Theo van Hintum, Centre for Genetic Resources, the Netherlands (CGN), Wageningen University & Research, Wageningen, the Netherlands.
Email: theo.vanhintum@wur.nl

Societal Impact Statement

As the world is facing a climate crisis and a growing population, feeding this population is a big challenge. Genebanks, conserving and providing access to genetic resources, and plant breeders, using genetic resources from genebanks to create new varieties, play important roles in meeting this challenge. Before making decisions potentially restricting access to digital sequence information (DSI) on genetic resources, it is therefore important to consider the impact of the decisions on the activities of these actors. In this paper, an analysis is made of DSI definitions and access and benefit-sharing scenarios in the context of their consequences for genebank management.

Summary

It is currently discussed whether the use of digital sequence information (DSI) on genetic resources would need to be subject to access and benefit-sharing obligations, like the use of genetic resources. In this paper, we analyse the consequences of genebank management of DSI definitions currently considered, and of scenarios proposed for dealing with the access and benefit-sharing aspects of DSI.

The analysis is based on publicly accessible literature and experiences of and discussions with genebank managers, researchers and experts on genetic resources policies.

The key findings are as follows:

- the definition of DSI is still disputed and definitions currently considered vary from only the base sequence of genomic DNA to all information associated with genetic resources;
- four groups of scenarios proposed for arranging the access and benefit-sharing aspects of DSI can be distinguished, with these scenarios differing in their benefit-sharing modalities, technical feasibility, the complexity they generate and thus the ease of access to information and genetic resources;
- from a genebank perspective, the scenario generating the lowest complexity and the easiest access to DSI would be preferable.

It is concluded that the multilateral and the free access scenarios seem most beneficial for genebanks, as these scenarios limit the complexity for users and allow easy access

and use. We are aware of the political difficulties to arrive at these solutions, but hope this paper will contribute to guiding the discussions in a direction that will be beneficial for genebanks, for users of genebank materials and information, and ultimately for addressing the challenges to present and future food security.

KEYWORDS
Access and Benefit Sharing (ABS), Convention on Biological Diversity (CBD), digital sequence information (DSI), genebanks, International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA), Nagoya Protocol

1 | INTRODUCTION

In the wake of increasing concerns on the loss of biodiversity, the Convention on Biological Diversity (CBD) was established in the early 1990s. The objectives of the CBD are as follows: (a) the conservation of biological diversity; (b) the sustainable use of its components and (c) the fair and equitable sharing of the benefits arising from the utilization of genetic resources (UNEP, 1992). The CBD came into force in December 1993. To further the third objective, a supplement to the CBD was drafted: the "Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization to the CBD." It came into force in October 2014, two decades after the CBD (UNEP, 2011). The CBD and the Nagoya Protocol prescribe the regulation of access to genetic resources and sharing of the benefits arising from their use in a bilateral way, on the basis of explicit permission (Prior Informed Consent, PIC) from the authorities of the provider country and a contract (Mutually Agreed Terms, MAT) between providers and users, necessitating case-by-case bilateral negotiations between users and providers. However, Parties to the Nagoya Protocol have the right to determine otherwise, and not to require PIC and MAT for access.

Both the CBD and the Nagoya Protocol deal with "genetic resources," which were defined by the CBD as "genetic material of actual or potential value" while "genetic material" was defined as "any material of plant, animal, microbial or other origin containing functional units of heredity" (UNEP, 1992). Additionally, international agreements with similar objectives have been set up for specific categories of genetic resources: the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA, in force since June 2004) for plant genetic resources for food and agriculture (PGRFA) and the Pandemic Influenza Preparedness Framework (PIP Framework, in force since May 2011) for influenza viruses with human pandemic potential. Different from the CBD and the Nagoya Protocol, the ITPGRFA and the PIP Framework are not based on bilateral agreements between providers and users, but have established multilateral systems for access and benefit-sharing, in which access is provided under Standard Material Transfer Agreements (SMTAs) instead of under the PIC and MAT terms prescribed by the CBD. In the case of the ITPGRFA, which is of particular importance for genebanks as it is focused on PGRFA, benefits are shared through a multilateral benefit-sharing fund used to support the conservation and sustainable utilization of PGRFA while the ITPGRFA also recognizes the benefit-sharing value of facilitated access to PGRFA, information exchange, capacity-building and access to and transfer of technology. In the PIP Framework, benefits are shared through preparedness and response projects in low-capacity countries while the Framework also facilitates access to vaccines and other countermeasures during a pandemic.

Since the coming into force of the CBD, technological developments in the biological and agricultural sectors have had a huge impact on the thinking about, and use of, genetic resources. More and more use is being made of genomic information next to or even instead of the genes themselves and (parts of) organisms containing the genes. As a result, the question was raised if the use of information on genetic resources, including information describing the DNA sequence of "functional units of heredity," would not also need to be subject to access and benefit-sharing obligations, like the use of genetic resources covered by the CBD, the Nagoya Protocol, the ITPGRFA and the PIP Framework. Presently, opinions diverge widely, and this issue has become a major barrier to finding agreement on other aspects of the international agreements.

The main fora where discussions on the access and benefit-sharing aspects of DSI take place are the CBD and the Nagoya Protocol, but DSI is also discussed in other fora, such as the ITPGRFA and the FAO Commission on Genetic Resources for Food and Agriculture (FAO-CGFRA). In the ITPGRFA, discussions on DSI started in 2017. In the latest Governing Board meeting (2019), discussions on DSI mainly took place within the framework of the negotiations on the enhancement of the Multilateral System of Access and Benefit-Sharing of the ITPGRFA. Mainly due to divergence of opinion on whether DSI should also be included in the ABS system of the ITPGRFA, these negotiations were not fruitful, and it was decided to await (and contribute to) DSI discussions and outcomes in the context of the CBD. The FAO-CGRFA established a work stream on "DSI on Genetic Resources for Food and Agriculture (GRFA)" in 2017, and commissioned a fact-finding scoping study to review the implications of the use of "DSI" for the conservation and sustainable use of GRFA, including exchange, access and the fair and equitable sharing of the benefits arising from their use. Further discussion was scheduled for the CGRFA meeting of 2020, but this meeting was postponed due to the COVID-19 pandemic.

Discussions in various fora have focused on the delimitation of the term DSI, with opinions ranging from it only comprising the sequence of nucleotides in DNA to all information related to genetic resources. Various alternatives for the term DSI have been used, reflecting this scope, including "genetic sequence data," "nucleotide

sequence data," "genetic sequences" or/and possibly most commonly, "digital sequence information" (DSI). The term "DSI" will be used in this paper as a placeholder.

Aside this definition issue, the discussions evolved around the central questions: should DSI be treated like genetic resources under the CBD, the Nagoya Protocol, the ITPGRFA and the PIP Framework and, if this would be the case, how benefit-sharing could be secured from the use of DSI.

Various studies have been conducted on the delimitation of DSI, and scenarios have been developed for the latter, ranging from business-as-usual to various bilateral and multilateral options. The aim of this paper is to examine the different DSI definitions and access and benefit-sharing scenarios proposed and to evaluate their possible implications for the day-to-day functioning of genebanks in their attempts to conserve plant genetic resources for future generations and make them accessible to the current.

2 | GENEBANKS AND DATA

Genebanks preserve genetic material and make it available for a wide range of users, including plant breeders, researchers, NGOs and farmers. As such, they play an important role in addressing the effects of climate change, population growth and other challenges to present and future food security (Brink & van Hintum, 2020). In this paper, the plant genebank of the Centre for Genetic Resources, The Netherlands (CGN) is taken as an example. Its information management is relatively complete, as it manages and shares all types of data in various ways. It may therefore not be a typical example of a genebank, but it serves to illustrate the "complete case." Other larger and more important genebanks, such as the International Genebanks managed in the CGIAR Genebank Platform, are largely in a similar situation, maintaining similar data and sharing these in similar ways.

2.1 | Types of data

Apart from the data used for logistics and distribution, such as the location of the seeds in the cold storage rooms or the details of a seed request, etc., the data in a genebank related to the genetic material comprise three distinct domains: passport, phenotype and various omics.

- Passport data describe the identity and origin of the material. These concern data on the taxonomic classification, population status (wild, landrace, cultivar, etc.), donor and identification numbers. Additionally, for collected material, there are data about the site where it was collected (including latitude, longitude, elevation), and for cultivated material, the breeder, ancestry and cultivar name. Each accession in the genebank collection will have one set of passport data that is completed and corrected as far as data are available, the amount of data will therefore be more or less constant in time.
- Phenotypic data describe the traits of the material. These concern (a) characterization traits: easy to observe, relatively independent of the environment and (b) evaluation traits: often dependent on the environment and requiring specific experiments or equipment to be determined. Flower colour or number of branches are examples of the first, sugar content or disease resistance examples of the second category. The number of data on phenotypic traits will increase linearly in time, as the material is regenerated more often and used in more experiments.
- Omics data come from high-throughput experiments generating very large amounts of data describing the genome (DNA), transcriptome (mRNA), proteome (proteins) or metabolome (substance involved in metabolism) in the plant cells. The various new data types generated in these experiments, such as the data generated by high-throughput phenotyping (phenomics), can also be included in this category. The number of omics data increases exponentially in time.

2.2 | Data management

In most genebanks, passport and phenotypic data are stored locally, in locally made applications of commercially available database management systems: genebank documentation systems. Various attempts have been made to create a generic genebank documentation system that can be adopted by various genebanks, the latest being GRIN-Global (Postman et al., 2010).

Omics data, currently generated at a large and increasing scale, are by default not included in the genebank documentation system but rather in dedicated other systems, either managed by the institute hosting the genebank or in public databases.

Obviously, much phenotypic and omics information, generated by the users of the genebank material in scientific or breeding experiments, is not made available to the genebank managing the germplasm. This information is stored in local databases, not accessible to the genebank or other users.

2.3 | Origin of genebank data

Passport data are collected by genebank staff prior to including a sample in the collection and validated and improved when appropriate. An example of an improvement of CGN's passport data is the inclusion of a "digital object identifier" (DOI), provided by the secretariat of the ITPGRFA, for every accession in the collection. This DOI will allow linking information in the local CGN documentation system to information in other databases, initially genomic data.

Characterization data, the simple phenotypic data, are generated by CGN staff during regenerations of the genebank material, based on fixed descriptor lists. Evaluation data, however, are generated by users of the material and made available to CGN. This concerns various types of experiments. There are joint projects, such as the ones funded by the EU, that include CGN material in their experiments and make the results available at some stage during or after the project. Other important sources of phenotypic data are the

collaborative screenings organized by CGN together with groups of users. In these screenings, the participants jointly test CGN material for a specific trait, often disease resistance. Typically, the generated data are shared by the participants, and CGN can make the data publicly available after an embargo of 3–5 years. The final source of phenotypic data is the regular users of CGN material, who are requested to share the non-confidential data they generated on the CGN material (on the basis of article 6.9 of the SMTA). However, this source is very limited, as strict enforcement of the duty to submit these data might result in fraudulent data entering the CGN databases.

Omics data relating to CGN material are always generated in externally funded projects, either by the user, where CGN agreed to make very large sets of genebank material available for analysis provided that the data would become available too, or in joint projects where CGN provided the material, and the data became publicly available afterwards.

2.4 | Data sharing

Most genebanks make their data publicly available. Local genebank documentation systems are often made accessible via the Internet, and data in the systems are often shared with online platforms that give access to information of various genebanks. The amount of data that are shared is generally rather limited though, usually restricted to passport data. Often, phenotypic data are not sufficiently organized to be made publicly available and omics data have not yet been generated. CGN is in a good position in this context as it has passport, phenotypic and omics data about its accessions, which are all publicly accessible.

The first way in which data are shared is through the genebank website. The data on the CGN website are refreshed every 2 months, and access is provided to all locally stored passport and phenotypic data, in various ways. All data are completely downloadable in Excel spreadsheets, most data are online searchable in a web-interface, and the passport data have been made machine readable based on semantic web technology (Finkers et al., 2015).

The second way of sharing data is by uploading them to aggregated databases. CGN uploads its passport data every 2 months to the European Search Catalogue for Plant Genetic Resources (EURISCO) and the Global Biodiversity Information Facility (GBIF), using a largely automatic procedure. EURISCO is the aggregated database containing passport data of most genebanks in Europe (Weise et al., 2017). GBIF is an international database that focuses on making data on any biodiversity, so not only crop diversity, available. CGNs phenotypic data have also been uploaded to EURISCO, but this is not yet a standard procedure. Data in EURISCO and GBIF are freely available and, for example, the EURISCO data are shared with Genesys, a global database with genebank data. Data from EURISCO and Genesys are also downloaded into the FAO's World Information and Early Warning System on Plant Genetic Resources for Food and Agriculture (WIEWS).

The third way in which CGN data are shared is via the Integrated Publishing Toolkit (ITP) provided by GBIF. This kit allows sharing data, in this case CGNs passport data, by placing them on the Internet in a highly standardized way using registered ontology terms. The data thus become completely FAIR (findable, accessible, interoperable and reusable) for other computers to use without human intervention (Wilkinson et al., 2016). CGN would like its phenotypic data to be included in this ITP dataset as well, but the necessary ontologies for the phenotypic data are not available yet.

For the omics data, CGN relies largely on the public databases. A small set of metabolomic data was made available on the genebank website, but no genomic data were published by CGN. CGN is still working on ways to improve the identification and linking of omics datasets related to its material. In that context, the introduction of DOIs for genebank material (Alercia et al., 2018) was very welcome, and CGN labelled all its material with DOIs. These easy to apply DOIs are being supplied on request without costs by the ITPGRFA Secretariat and serve as permanent unique identifiers for genebank material. They allow automatic linking of information from different sources and thus facilitate storage of information on genebank accessions in different specialized databases. The use of DOIs will not only greatly increase access to information but will also allow tracking of the associated accessions: making visible where the germplasm was collected, in which genebanks it was stored and how it was used in research and breeding.

3 | DEFINITION OF DSI

The term "DSI" is being used as a placeholder, and so far, no agreement has been reached on what exactly it comprises (Laird et al., 2020). In its 2018 meeting, the Ad Hoc Technical Expert Group (AHTEG) on DSI on Genetic Resources established under CBD and its Nagoya Protocol proposed the following list of types of information that could be included in the term (CBD, 2018):

a. The nucleic acid sequence reads and the associated data;
b. Information on the sequence assembly, its annotation and genetic mapping;
c. Information on gene expression;
d. Data on macromolecules and cellular metabolites;
e. Information on ecological relationships and abiotic factors of the environment;
f. Function, such as behavioural data;
g. Structure, including morphological data and phenotype;
h. Information related to taxonomy;
i. Modalities of use.

However, within the AHTEG, opinions differed on which of these types of information should be included in DSI, from limiting DSI to the first group only to including all nine types of information.

Later, in a study commissioned by the CBD, Houssen et al. (2020) brought delineation of DSI down to four possible cumulative groups of information:

1. Narrow: DNA and RNA
2. Intermediate: DNA, RNA and proteins
3. Intermediate: DNA, RNA, proteins and metabolites
4. Broad: DNA, RNA, protein, metabolites and traditional knowledge, ecological interactions, etc.

In its 2020 meeting, the AHTEG considered these four groups and agreed that the first three groups could be considered as DSI while associated information included in fourth and broadest group would not be DSI (CBD, 2020). The outcomes of the AHTEG will be discussed in the third meeting of the Open-ended Working Group (OEWG) on the Post-2020 Global Biodiversity Framework (originally planned to be held in 2020, but postponed to 2021 due to the COVID-19 pandemic), and the OEWG will make recommendations to the CBD Conference of the Parties at its fifteenth meeting (COP15), also postponed to 2021.

In this respect, it needs to be noted that, whatever the COP15 will conclude, various countries already have made their own interpretations, some of which very broad, and incorporated these in their national Access and Benefit-Sharing (ABS) legislations. The ABS law of Malaysia, for instance, covers "biological resources," including "any information relating to" genetic resources, populations and biotic components (Lawson et al., 2019).

4 | ACCESS AND BENEFIT-SHARING SCENARIOS

A variety of scenarios to deal with DSI on genetic resources have been proposed. The proposed scenarios in publicly available papers can be roughly divided into four groups (Table 1):

4.1 | S1. Current situation

In the current situation, "business as usual," some countries consider DSI on genetic resources to be in scope of the Nagoya Protocol, either by mentioning the inclusion of DSI in their legislation or by stating that DSI is to be considered a genetic resource while others do not consider DSI to be in scope. If a country decides to include DSI in its access and benefit-sharing legislation, access to and use of DSI from this country may not be free anymore, as potential users are bound by the national laws of the country. Already more than 15 countries are currently including DSI in their access and benefit-sharing legislation (Bagley et al., 2020), in a wide diversity of ways, resulting in considerable complexity.

Also, provider countries of genetic resources may want to have included provisions in the MAT associated with access to these genetic resources stipulating that any DSI derived from these genetic resources cannot be made publicly available without permission from the country providing the genetic resources. This means that various types of information would not be available for use without permission of the country providing the genetic resources to which the DSI applies. DSI on genetic resources uploaded in databases would need to get the so-called country tags, to enable the user to obtain PIC and/or MAT from the provider country of the genetic resources to which the DSI applies, if required by that country.

4.2 | S2. Bilateral access and benefit-sharing systems

In the most obvious variant of this scenario, DSI would be considered equivalent to genetic resources, and domestic access and

TABLE 1 DSI scenarios/options presented in three published papers

Source	1. Current situation (DSI not explicit in scope NP, but may be included in national legislation and in MAT)	2. Bilateral access and benefit-sharing systems	3. Multilateral access and benefit-sharing systems	4. Free access to DSI
Hiemstra et al. (2019)	Scenario 1 (DSI out of scope of the NP)	Scenario 2 (DSI equivalent to genetic resources; no additional measures)	Scenario 3 (DSI out of scope of the NP, but multilateral benefit-sharing)	Scenario 4 (free exchange within coalition of willing)
First Global Dialogue on DSI (Anon., 2020)		Option 1 (Nagoya—bilateral benefit-sharing) Option 2 (open access, bilateral benefit-sharing for commercial use through country tag)	Option 3 (open access; benefit-sharing in case of commercial use; multilateral fund) Option 4 (open access; benefit-sharing through subscription fee/levies; multilateral fund)	Option 5 (free access with capacity development)
Scholz et al. (2020)		Option 4 (commons licenses for DSI) Option 5 (blockchain metadata, open DSI)	Option 1 (micro-levy) Option 2 (membership fee) Option 3 (cloud-based fees)	Option 0 ("status quo": DSI & non-monetary benefit-sharing)[a]

Abbreviations: DSI, digital sequence information; MAT, Mutually Agreed Terms; NP, Nagoya Protocol.

[a]Although this option is named "status quo," its description seems to imply free access.

benefit-sharing legislation for genetic resources would automatically apply to DSI as well (Hiemstra et al., 2019). This means that PIC and MAT for DSI would be required if they are also required for genetic resources, unless countries explicitly exempt DSI from their access and benefit-sharing rules.

However, DNA sequence data (which uncontestably form part of DSI) are generally not used individually, but derive their value through comparison with, often, many other sequence data (Laird & Wynberg, 2018). Therefore, the current system of the Nagoya Protocol (PIC and MAT for one or a group of genetic resources from one country) is often not considered suitable to be applied for sequence data, because it would mean that for a comparison of sequences from different countries PIC and MAT will have to be sought from all these countries. Furthermore, the precise contribution of each separate piece of DSI to the end product, and thus the required amount of benefit-sharing connected with its use, will be impossible to determine.

In view of this, alternative bilateral access and benefit-sharing systems for DSI have been proposed which do not make use of PIC and MAT, strive for open access, but would still ensure that benefits are shared with the specific provider countries of the genetic resources to which the DSI applies. For these bilateral access and benefit-sharing systems, a country tag should be connected to DSI in databases, and users of this DSI could be informed of the obligation to share benefits with the provider countries of the genetic resources to which the DSI applies, for instance when the DSI is used for commercial applications (Anon, 2020). Proposed variants of this scenario make use of blockchain technology to track the use of DSI or commons licenses that set out the terms of use of the DSI (Scholz et al., 2020).

4.3 | S3. Multilateral access and benefit-sharing systems

Bilateral access and benefit-sharing systems require tracking and tracing to determine the origin of the genetic resources to which the DSI applies and the utilization of that DSI, which may necessitate complex mechanisms. In multilateral access and benefit-sharing systems, this would be avoided, as access and benefit-sharing would be decoupled. This means that benefits arising from the use of DSI are not directly shared with the provider countries of the genetic resources from which the DSI was derived, but that these benefits are shared through a multilateral fund.

Multilateral solutions also solve the problem that often many sequences are compared and that the precise contribution of each separate piece of DSI to the end product and thus the required amount of benefit-sharing connected with its use are often impossible to determine. This resembles the problem in assessing the value of each separate PGRFA used in breeding new cultivars, where often genetic material from many different PGRFAs is combined (Gaffney et al., 2020). This was one of the reasons for developing the multilateral system of the ITPGRFA.

In this scenario, DSI accessed under the multilateral access and benefit-sharing system would not fall under the provisions of the Nagoya Protocol, in a similar way that genetic resources accessed in the framework of the multilateral system of the ITPGRFA are not in scope of the Nagoya Protocol. For this, the multilateral benefit-sharing system for DSI could be recognized under the Nagoya Protocol as a specialized international instrument, like the ITPGRFA. Another possibility would be to differentiate between types of DSI (e.g. DSI from PGRFA, DSI from human pathogens), and to include DSI on genetic resources from specific categories of genetic resources in specialized instruments for these genetic resources, such as the already existing ITPGRA and PIP Framework.

Various variants of this scenario have been proposed, including obligatory benefit-sharing when a product based on accessed DSI is commercialized, subscription systems, and levies on equipment used in sequencing (Anon, 2020; Aubry, 2019; Lawson et al., 2019; Scholz et al., 2020). Although in these variants conditions will be attached to the access and utilization of DSI, access basically remains open for members of the multilateral system.

It has also been suggested to establish an opt-in possibility to also include genetic resources in a multilateral benefit-sharing system for DSI (Scholz et al., 2020). A scenario with DSI as well as genetic resources covered in one multilateral system with open, but not free, access, would reflect the concept of "bounded openness," proposed by Vogel et al. (2011).

4.4 | S4. Free access

In this scenario, in contrast to the preceding ones, it is agreed that no specific monetary benefit-sharing obligations are connected to the access to or utilization of DSI, and access to DSI is free. Arguments brought forward in favour of this scenario are that free availability of information in itself is a form of benefit-sharing, and that the use of DSI leads to various forms of indirect benefit-sharing, such as the availability of improved plant varieties (Gaffney et al., 2020). To increase this type of indirect benefit-sharing, increased cooperation and capacity building in low- and middle-income countries is advocated (Gaffney et al., 2020).

In a variant of this option, there may be no universal-free access to DSI, but countries may form a "coalition of the willing" with free exchange of DSI (and perhaps also genetic resources) among its members, and restricted access to others (Hiemstra et al., 2019).

5 | IMPLICATIONS OF DSI DEFINITIONS AND ACCESS AND BENEFIT-SHARING SCENARIOS FOR THE DAY-TO-DAY FUNCTIONING OF GENEBANKS

CGN aims to make the accessions in its collection easily available for research, breeding or training for food and agriculture purposes, using the SMTA of the International ITPGRFA. Likewise, it aims to make the

information it has collected on these accessions (passport data, phenotypic data and/or omics data) easily available to users, to enhance the effectiveness of further research and breeding. Regulation of access to and use of information could mean that CGN would no longer be able to provide this information without restrictions on its further use. To what extent this would be the case depends on the definition of DSI and the type of regulation as reflected in the various scenarios above.

5.1 | Definition of DSI

When the three types of data related to the genetic material in genebanks as distinguished in Section 2 (passport data, phenotypic data and omics data) are compared with the proposed types of information that could be included in the term DSI (Section 4), it is clear that the wider the definition of DSI, the more genebank data will be affected.

Passport and phenotypic data would only be included in DSI when very broad definitions would be agreed upon: in CBD (2018) (e), (f) and (g) for phenotypic data and (i) and perhaps (h) for passport data, and for both category (4) in Houssen et al. (2020). Omics data, on the other hand, would also fit in more narrow definitions of DSI [(a), (b), (c) and (d) in CBD (2018), and (1), (2) and (3) in Houssen et al. (2020)], depending on the types of omics data.

It is hard to imagine how genebank documentation systems could remain publicly accessible and usable if broader definitions of DSI would be agreed upon, and the publication and utilization of passport data and phenotypic data would be subject to access and benefit-sharing regulations.

Passport data and phenotypic data are essential for genebank management: improving the composition of the collection and monitoring of the genetic integrity during regenerations. But obviously they are also the data types that identify the germplasm and serve as a first lead or selecting material for use. Therefore, the availability of this type of information increases the value of the genetic resources, as it enables users to better select the material that meets the requirements and has the desired properties. Thus, the acceptance of broader definitions of DSI would strongly reduce the value of the material in the genebank, and its utilization.

The use of the narrower definitions of DSI, coupled with a restricted access to this information, would have less direct impact on genebanks in the short term, as the use of this type of information in genebanks is still very restricted. However, as we speak, much PGR held in genebanks is being sequenced, and the possibilities of using this new data are being explored to improve the services of genebanks: by improving the composition of the collection and improving the efficiency of selecting material with the desired traits.

5.2 | Access and benefit-sharing scenarios

What would the four scenarios mean for genebanks? In the discussion of the scenarios, it is assumed that the definition of DSI is narrow or intermediate. If, on the other hand, access to and utilization of passport and characterization data are restricted, the functioning of genebanks will be seriously impeded, as explained above.

5.2.1 | S1. Current situation

In the current situation, access to DSI is only regulated in countries which include DSI in their domestic rules (either by mentioning the inclusion of DSI in their legislation or by stating that DSI is to be considered a genetic resource). If no international agreement is reached on how to deal with the access and benefit-sharing aspects of DSI, it can be expected that more and more countries will include DSI in their domestic access and benefit-sharing legislation. They will do this in a diversity of ways, resulting in much complexity. It would imply that the conditions for sharing information about genebank accessions would differ per accession. This could lead to genebanks deciding not to include material from certain countries in their collections but to avoid the complexity for both themselves and their users, in using the associated information.

This is very similar to what has happened with genetic resources. Due to the current complexity of domestic regulation of access to genetic resources, following the establishment of the CBD and the Nagoya Protocol, the international exchange of PGRFA as well as their utilization has been seriously hampered, and benefit-sharing has fallen short of expectations (Brink & van Hintum, 2020; Laird et al., 2020). In this respect, it must be remarked that lack of access to genetic resources (and possibly DSI) of a certain country does not only affect potential users but also the country blocking access, as it may not only lead to lack of benefit-sharing but also to isolation of this country, as its genetic resources may to a lesser extent be subject to scientific research and its researchers may be less involved in international scientific cooperation.

In situations where countries require the inclusion of provisions on DSI in the MAT associated with access to the genetic resources from which DSI is derived, publication of this DSI may be prohibited without consent from the provider country. If passport data and phenotypic data would be included in the definition of DSI, this would mean that genebanks would have to seek permission from provider countries to be able to publish even this basic information.

A final consideration is that CGN and other genebanks use the SMTA of the ITPGRFA to distribute the material in their collections. As this SMTA only concerns genetic resources and do not cover information, countries may become more reluctant to make their genetic resources available to genebanks under the SMTA.

For all these reasons, maintaining the status quo (Scenario 1) thus seems not to be a favourable option for genebanks.

5.2.2 | S2. Bilateral access and benefit-sharing systems

If it would be simply agreed that DSI is equivalent to genetic resources and that PIC and MAT are required for access to and use

of DSI (where required for genetic resources), the result would be increased complexity and decreased access, like explained above for Scenario 1, or even more, which would clearly be an unfavourable development for genebanks.

Other variants of bilateral access and benefit-sharing systems for DSI which do not require PIC and MAT, but, for instance, entail the obligation to share benefits with the provider countries of the genetic resources to which the DSI applies when the DSI is used for commercial applications, could be less burdensome for genebanks. Especially when bilateral benefit-sharing on DSI would not be coupled to access, but to utilization (for any purpose or only for commercial purposes), genebanks, which usually only make material and information available, without utilization by the genebanks themselves, would be able to publish DSI without having to seek permission from all provider countries of the genetic resources to which the DSI applies. Of course, genebanks will need to attach country tags to the DSI they make available, to enable third-party users to fulfil their benefit-sharing obligations. However, as it is already standard practice for genebanks to include data on the origin of their genetic resources in the passport data and as the DSI provided by genebanks will normally only be related to genetic resources in these genebanks, provider countries of the genetic resources to which the DSI applies can be identified easily.

In practice, especially if the PIC and MAT system would be applicable to DSI, it could be expected that users of DSI would concentrate more on DSI from material collected prior to the moment that the PIC and MAT system would apply for DSI (although various countries have included elements of retroactivity in their domestic access legislation). Genebanks might respond by expanding their collections of genetic resources and associated information by merely exchanging "old material" and information between genebanks. Collecting new material would become more difficult and unattractive, similar effects as those of the introduction of domestic ABS legislation based on the CBD and the associated Nagoya Protocol (Brink & van Hintum, 2020). The resulting complexity and practical difficulties discouraged or even inhibited further collecting of valuable genetic resources, thus hindering the prevention of the genetic erosion that is occurring due to the climate crisis and complicating the breeding activities needed to create the varieties needed to feed the growing world population.

Another effect of restricting access to, and thus use of, DSI is that the production and management of this information will increasingly move from the public domain to the private. This is also similar to an effect that could be observed with regard to genetic resources prior to the introduction of the Nagoya Protocol, when private companies built up their own genebank collections to become less dependent of public genebanks. This shift from public to private is obviously undesirable, as it will reduce the volume of data publicly available for both public and private research.

An important additional consideration is whether regulation of access to and benefit-sharing from DSI would only apply to DSI acquired after entry into force of this regulation (new acquisitions), or if it would also apply to DSI acquired earlier, for example, all DSI acquired since the entry into force of the CBD. In the case of regulation of access to genetic resources through the Nagoya Protocol (which came into force on 12 October 2014), most countries consider only genetic resources acquired on or after that date to fall in scope of the Nagoya Protocol while other countries also consider material acquired earlier than that date to fall under the Nagoya Protocol (depending on the entry into force and the provisions of domestic ABS legislation). A special case is Brazil, which defines "access" as "research on and with genetic heritage," which means that "access" can be at a (much) later date than acquisition. If regulation of access to DSI would also cover DSI acquired from or entered into public databases before its entry into force, the possibilities for genebanks to make information available would be further reduced than when only DSI acquired after entry into force would be affected.

5.2.3 | S3. Multilateral access and benefit-sharing systems

Multilateral access and benefit-sharing systems promise to be less complex for users than bilateral access and benefit-sharing systems (especially the Nagoya system based on PIC and MAT). Therefore, the implementation of multilateral access and benefit-sharing systems for DSI would be a better option for genebanks, bringing less technical difficulties and transaction costs. However, some complexity may remain if distinctions have to be made between users which have subscribed to the multilateral system and those who have not.

Of course, complexity would be even further reduced if genetic resources would also be included in a multilateral benefit-sharing system for DSI, as suggested by Scholz et al. (2020). Reduction of complexity of access to genetic resources is important, as increasing doubt has arisen about the bilateral approach of the Nagoya Protocol, because it makes academic and conservation research much more difficult while not generating substantial benefits for biodiversity conservation (Laird et al., 2020).

For PGRFA, a multilateral solution for the DSI problem could be partially implemented by incorporating DSI on PGRFA in the SMTA of the ITPGRFA, but this would not solve the situation for DSI from PGRFA which are not mentioned in Annex I of the ITPGRFA and thus not included in the Multilateral System of the ITPGRFA, and for uses beyond those covered by the SMTA (training, research and breeding for food and agriculture) (Aubry, 2019).

Also for multilateral ABS systems for DSI, it is important to know if regulation of access to and benefit-sharing from DSI would only apply to DSI acquired after entry into force of this regulation or if it would also apply to DSI acquired earlier, as set out under S2.

5.2.4 | S4. Free access

In the short run, this scenario would have very limited or no consequences for genebank operations. In a "coalition of the willing" scenario, where the members of this coalition would have free access to

all genetic resources and associated information managed within the coalition, the only issue, from a genebank perspective, would be to make the distinction between requests that are part of the system and those that are not.

It is hard to imagine that this option, which does not involve direct monetary benefit-sharing but does recognize non-monetary benefit-sharing, would be acceptable for all countries. It would not contribute to building trust between countries, and could endanger agreement on other CBD, Nagoya Protocol, ITPGRFA and PIP Framework issues as well. In the long run, it may therefore have detrimental effects on the availability of genetic resources as well as on non-monetary benefit-sharing.

On the other hand, free exchange of genetic resources and associated information can also be considered to be beneficial to all parties involved given the close mutual dependence (Khoury et al., 2016) and the vital importance of active exchange and use of these resources to counteract the effects of climate change and population growth.

6 | CONCLUSIONS AND RECOMMENDATIONS

Having the possibility to make available genetic and other information on the genetic resources in their collection is important for genebanks to be able to make available the resources needed to meet the demands of a growing world population in a changing climate. Restriction of this possibility to make information available through access and benefit-sharing regulations on DSI may jeopardize the functioning of genebanks. Therefore, the effects in practice of various proposals concerning the definition of DSI and access and benefit-sharing scenarios should be carefully evaluated.

As for the definition of DSI, extending this definition beyond genetic sequence data, that is, beyond category (d) of CBD (2018) or group 3 of Houssen et al. (2020), may seriously hamper genebanks in making available passport and phenotypic data, with undesirable consequences for genebank management and access to the PGRFA in the genebanks. The use of narrower definitions of DSI, coupled with restricted access to this information, would have less direct impact on genebanks in the short term.

As for the scenarios proposed for arranging the access and benefit-sharing aspects of DSI, Scenario 4 (free access), although politically difficult to achieve, would be the most attractive from a genebank perspective, as it minimizes complexity and avoids privatization of DSI. Furthermore, full public access to as much DSI as possible will allow optimal use of PGR to counteract the consequences of the imminent threats the world is currently facing.

Continuation of the current situation (Scenario 1) is expected to lead to increased complexity, as more and more countries are expected to include DSI in their domestic access and benefit-sharing legislation, in a diversity of ways. This increased complexity will hinder the functioning of genebanks and may lead to genebanks no longer conserving material from certain countries. This will not only affect potential users but may also affect the countries regulating access to DSI, as their genetic resources may to a lesser extent be subject to scientific research and its researchers may be less involved in international scientific cooperation. The effects of this increased complexity on genebank functioning will be especially severe in case it is opted for a wide definition of DSI.

Agreement on bilateral access and benefit-sharing systems (Scenario 2) also brings the risk of increased complexity and decreased access. This would certainly be the case if it would be agreed that DSI is equivalent to genetic resources and that PIC and MAT are required for access to and use of DSI, where they are required for genetic resources. Other variants of bilateral access and benefit-sharing systems for DSI, not requiring PIC and MAT, could be less burdensome for genebanks, especially when bilateral benefit-sharing on DSI would not be coupled to access, but to utilization. Like for scenario 1, the effects of this increased complexity on genebank functioning will be more severe in case it is opted for a wide definition of DSI.

Multilateral access and benefit-sharing systems promise to be less complex, as they do not require elaborate tracking and tracing systems or even necessitate case-by-case bilateral agreements between users and providers for each unit of DSI. They seem thus more favourable for genebank functioning than bilateral access and benefit-sharing systems. For PGRFA, a multilateral solution for the DSI problem solution could be partially implemented by incorporating DSI in the SMTA of the ITPGRFA, but this would not solve the situation for DSI from PGRFA which are not mentioned in Annex I of the ITPGRFA and for uses beyond those covered by the SMTA.

For bilateral as well as for multilateral ABS systems for DSI, it is important to know if regulation of access to and benefit-sharing from DSI would only apply to DSI acquired after entry into force of this regulation or if it would also apply to DSI acquired earlier. If regulation of access to DSI would also cover DSI acquired from or entered into public databases before its entry into force, the possibilities for genebanks to make information available would be further reduced than when only DSI acquired after entry into force would be affected.

We hope this paper will contribute to guiding the discussions in a direction that will be beneficial for genebanks, for users of genebank materials and information, and ultimately for addressing the challenges to present and future food security.

ACKNOWLEDGEMENTS

The authors would like to thank the constructive feedback from the editor and reviewers, which greatly helped to improve the quality of the manuscript.

AUTHOR CONTRIBUTIONS

MB and TvH planned and designed the article. MB wrote the first draft of Sections 1, 3, 4 and 6, and TvH wrote the first draft of the abstract and Sections 2 and 5. Both authors contributed to manuscript revision, and read and approved the submitted and final versions.

REFERENCES

Alercia, A., López, F. M., Sackville Hamilton, N. R., & Marsella, M. (2018). *Digital Object Identifiers for food crops—Descriptors and guidelines of the Global Information System*. FAO.

Anon. (2020). *Report first global dialogue on digital sequence information on genetic resources, 6–8 November 2019, Pretoria, South Africa*. The ABS Capacity Development Initiative, the South African National Department of Environment, Forestry and Fisheries, and the Norwegian Government.

Aubry, S. (2019). The future of digital sequence information for plant genetic resources for food and agriculture. *Frontiers in Plant Science*, *10*. https://doi.org/10.3389/fpls.2019.01046

Bagley, M., Karger, E., Ruiz Muller, M., Perron-Welch, F., & Thambisetty, S. (2020). *Fact-finding study on how domestic measures address benefit-sharing arising from commercial and non-commercial use of digital sequence information on genetic resources and address the use of digital sequence information on genetic resources for research and development*. Convention on Biological Diversity Montreal.

Brink, M., & van Hintum, T. (2020). Genebank operation in the arena of access and benefit-sharing policies front. *Plant Science*, *10*. https://doi.org/10.3389/fpls.2019.01712

CBD. (2018). *Report of the Ad Hoc Technical Expert Group on digital sequence information on genetic resources*. Convention on Biological Diversity, Montreal, Canada, 13–16 February 2018, 10 pp.

CBD. (2020). *Report of the Ad Hoc Technical Expert Group on digital sequence information on genetic resources*. Convention on Biological Diversity, Montreal, Canada, 17–20 March 2020, 18 pp.

Finkers, H. J., Chibon, P. Y. F. R. P., van Treuren, R., Visser, R. G. F., & van Hintum, T. J. L. (2015). Genebanks and genomics: How to interconnect data from both communities? *Plant Genetic Resources: Characterization and Utilization*, *13*, 90–93. https://doi.org/10.1017/S1479262114000689

Gaffney, J., Tibebu, R., Bart, R., Beyene, G., Girma, D., Kane, N. A., Mace, E. S., Mockler, T., Nickson, T. E., Taylor, N., & Zastrow-Hayes, G. (2020). Open access to genetic sequence data maximizes value to scientists, farmers, and society. *Global Food Security*, *26*, 100411. https://doi.org/10.1016/j.gfs.2020.100411

Hiemstra, S. J., Brink, M., & van Hintum, T. J. L. (2019). *Digital Sequence Information (DSI): Options and impact of regulating access and benefit sharing—Stakeholder perspectives*. CGN report 42. Centre for Genetic Resources, the Netherlands (CGN), Wageningen, The Netherlands, 22 pp.

Houssen, W., Sara, R., & Jaspars, M. (2020). *Digital sequence information on genetic resources: Concept, scope and current use*. Convention on Biological Diversity Montreal.

Khoury, C. K., Achicanoy, H. A., Bjorkman, A. D., Navarro-Racines, C., Guarino, L., Flores-Palacios, X., Engels, J. M. M., Wiersema, J. H., Dempewolf, H., Sotelo, S., Ramírez-Villegas, J., Castañeda-Álvarez, N. P., Fowler, C., Jarvis, A., Rieseberg, L. H., & Struik, P. C. (2016). Origins of food crops connect countries worldwide. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1832), 20160792. https://doi.org/10.1098/rspb.2016.0792

Laird, S. A., & Wynberg, R. P. (2018). *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*. Convention on Biological Diversity Montreal.

Laird, S., Wynberg, R., Rourke, M., Humphries, F., Ruiz Muller, M., & Lawson, C. (2020). Rethink the expansion of access and benefit sharing. *Science*, *367*(6483), 1200–1202. https://doi.org/10.1126/science.aba9609

Lawson, C., Humphries, F., & Rourke, M. (2019). The future of information under the CBD, Nagoya Protocol, Plant Treaty, and PIP Framework. *Journal of World Intellectual Property*, *22*(3–4), 103–119. https://doi.org/10.1111/jwip.12118

Postman, J., Hummer, K., Ayala-Silva, T., Bretting, P., Franko, T., Kinard, G., Bohning, M., Emberland, G., Sinnott, Q., Mackay, M., Cyr, P., Millard, M., Gardner, C., Guarino, L., & Weaver, B. (2010). GRIN-global: An international project to develop a global plant genebank information management system. *Acta Horticulturae*, *859*, 49–55. https://doi.org/10.17660/ActaHortic.2010.859.4

Scholz, A., Hillebrand, U., Freitag, J., Cancio, I., Ribeiro, C., Haringhuizen, G., Oldham, P., Saxena, D., Seitz, C., Thiele, T., & van Zimmeren, E. (2020). *Finding compromise on ABS & DSI in the CBD: Requirements & policy ideas from a scientific perspective*. Wissensbasierte Lösungsansätze für Digitale Sequenzinformation (WiLDSI), 42 pp.

UNEP. (1992). *Convention on biological diversity. Text and annexes*. United Nations Environmental Programme (UNEP), Secretariat of the Convention on Biological Diversity.

UNEP. (2011). *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. Text and annex*. United Nations Environmental Programme (UNEP), Secretariat of the Convention on Biological Diversity.

Vogel, J. H., Álvarez-Berríos, N., Quiñones-Vilches, N., Medina-Muñiz, J. L., Pérez-Montes, D., Arocho-Montes, A., Val-Merniz, N., Fuentes-Ramírez, R., Marrero-Girona, G., Mercado, E. V. & Santiago-Ríos, J. (2011). The economics of information, studiously ignored in the Nagoya Protocol on access to genetic resources and benefit sharing. *Law, Environment and Development Journal*, *7*(1), 52–65.

Weise, S., Oppermann, M., Maggioni, L., van Hintum, T., & Knüpffer, H. (2017). EURISCO: The European search catalogue for plant genetic resources. *Nucleic Acids Research*, *45*(1), D1003–D1008. https://doi.org/10.1093/nar/gkw755

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, *3*, 160018. https://doi.org/10.1038/sdata.2016.18

How to cite this article: Brink M, van Hintum T. Practical consequences of digital sequence information (DSI) definitions and access and benefit-sharing scenarios from a plant genebank's perspective. *Plants, People, Planet*. 2021;00:1–10. https://doi.org/10.1002/ppp3.10201