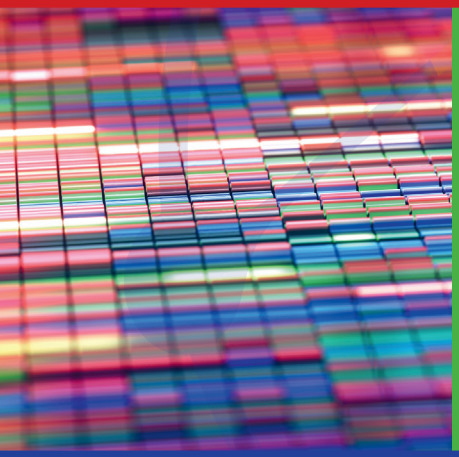
THE ABS CAPACITY DEVELOPMENT INITIATIVE

L'INITIATIVE DE RENFORCEMENT DES CAPACITES POUR L'APA



Digital Sequence Information on Genetic Resources (DSI)

An Introductory Guide for African Policymakers and Stakeholders

Authors:
Elizabeth Karger
Pierre du Plessis
Hartmut Meyer

www.abs-initiative.info

The ABS Initiative is funded by

Federal Ministry for Economic Cooperation and Development

Organisation Internationale de la Francophonie

Norwegian Ministry of Foreign Affairs

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

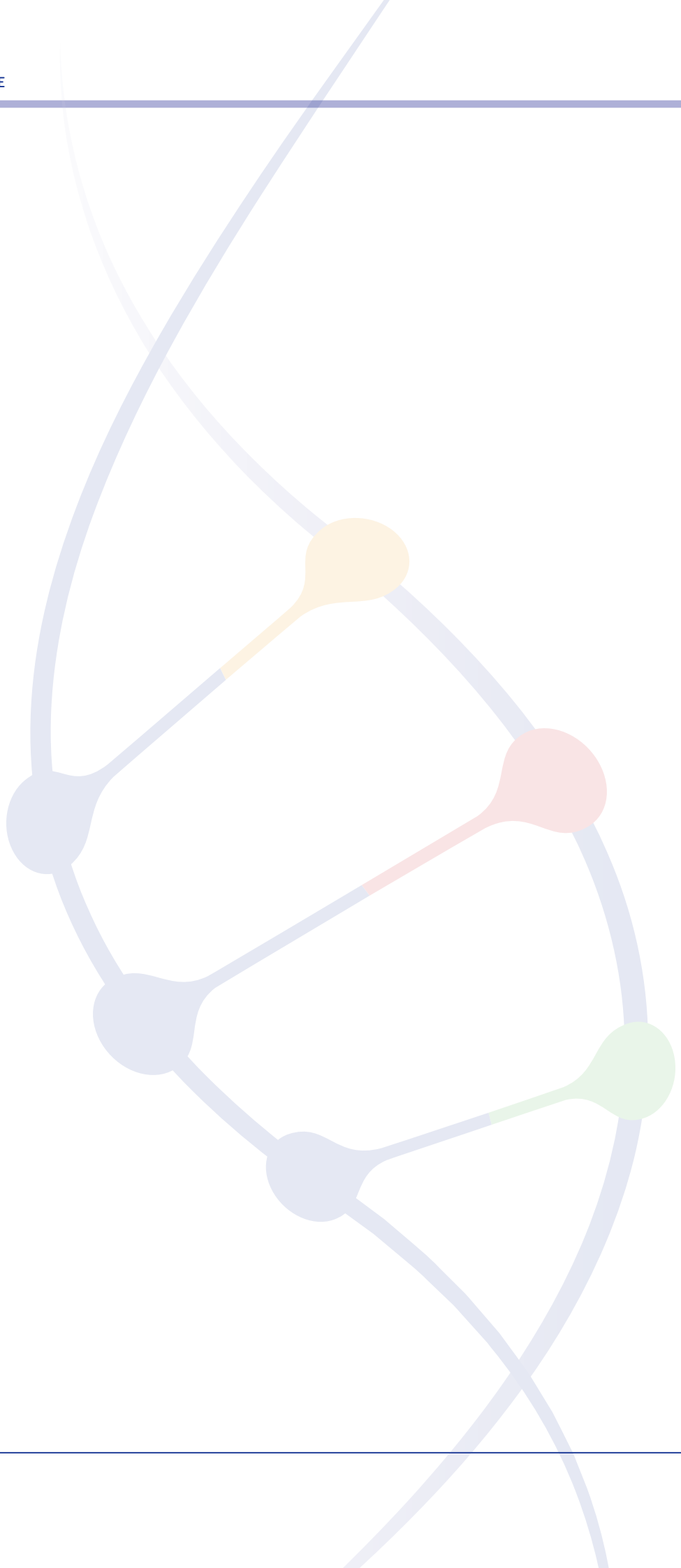
Swiss Confederation

Federal Department of Economic Affairs, Education and Research EAER
State Secretariat for Economic Affairs SECO

Institut de la Francophonie pour le Développement Durable IFDD

and implemented by

giz Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH



EXECUTIVE SUMMARY

Digital sequence information on genetic resources (DSI) has had a dedicated work stream on the international agenda of the *Convention on Biological Diversity* (CBD) and its *Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization* (Nagoya Protocol) since 2016.

The term DSI does not have an internationally agreed meaning and is not commonly used among scientists. It was first introduced into discussions under the CBD and has since become a placeholder, although its concept and scope are not clear and more appropriate terminology is needed.

The focus of the international discussion on DSI has so far been on sequential molecules, such as deoxyribonucleic acid (DNA), ribonucleic acid (RNA) and proteins (amino acids), but it is possible that other types of information are also relevant to the discussion.

The issue of DSI arose due to the increasing speed and falling costs of sequencing, which have resulted in an enormous quantity of biological data being produced and stored in publicly accessible databanks, which are used for research and development, including for commercial purposes. This happens in the absence of benefit-sharing obligations and many Parties and other actors are concerned that this will negatively impact on the third objective of the CBD and the objective of the Nagoya Protocol.

Divergent positions on DSI have emerged over the past few years. There is disagreement, for example, on whether the definition of "genetic resources" in the CBD covers DSI or not, whether DSI should fall under the access and benefit-sharing (ABS) regime, and whether open access to DSI can be regarded as a sufficient form of benefit-sharing. These issues are being dealt with in the context of the CBD and other international instruments and processes.

At the CBD's 14th Conference of the Parties (COP) in Sharm El Sheik, Egypt in November 2018, key decisions were made on DSI (Decision14/20) and the comprehensive and participatory process for the preparation of the post-2020 global biodiversity framework (Post- 2020 Framework) (Decision 14/34). Decision 14/20 put in place a science policy process to inform the discussions on DSI. DSI has also been plugged into the intersessional activities supporting the development of the Post-2020 Framework, making the Framework central to how DSI will be dealt with in the future. Central questions for African actors include how the Post-2020 Framework will address the three objectives of the CBD in a balanced way, what role ABS should play, what the link is between ABS, DSI and the Sustainable Development Goals (SDGs) and critically how DSI, and especially benefit-sharing, will be addressed.

The Parties to the CBD and the Nagoya Protocol will meet again in 2020 in Kunming, China, at which time further decisions will be made on DSI and the Post-2020 Framework. A number of Parties, including the African Group, have indicated that benefit-sharing is an inextricable part of the CBD "package deal" and the Post-2020 Framework must include benefit-sharing for DSI if the Post-2020 Framework is to be accepted.

A NOTE ON TERMINOLOGY

The term **"digital sequence information on genetic resources"**, commonly referred to as **DSI**, does not have an internationally agreed meaning and is not commonly used among scientists. It is a term that was first introduced into discussions under the CBD and its Nagoya Protocol and has become a placeholder, although its concept and scope are not clear.

There are various other terms that might be used in practice, including "genetic sequence data", "genetic sequence information", "genetic information", "dematerialized genetic resources", "in silico utilization", etc.

In January 2018, the CBD's Ad Hoc Technical Expert Group (AHTEG) on DSI met to consider, among other things, the technical scope and legal and scientific implications of existing terminology related to DSI. The experts:

- identified various types of information that may be relevant to the utilization of genetic resources and the objectives of the CBD and its Nagoya Protocol, including, among other things:
  - → nucleic acid sequence reads and associated data;
  - → information on the sequence assembly, its annotation and genetic mapping. This information may describe whole genomes, individual genes or fragments thereof, barcodes, organelle genomes or single nucleotide polymorphisms;
  - → information on gene expression;
  - → data on macromolecules and cellular metabolites;
  - → information on ecological relationships, and abiotic factors of the environment;
  - → function, such as behavioural data;
  - → structure, including morphological data and phenotype;
  - → information related to taxonomy; and
  - → modalities of use;
- reached consensus that DSI is not the appropriate term to refer to these types of information but would continue to be used as a placeholder.

In this Introductory Guide, the placeholder term "DSI" will be used.

INTRODUCTION

Digital sequence information on genetic resources (DSI) found its way onto the international agenda of the *Convention on Biological Diversity* (CBD) and its *Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization* (Nagoya Protocol) in 2016. Since then, there has been a great deal of discussion on DSI at the international level, but not much progress has been made towards reaching consensus on how to deal with this issue. The Parties to the CBD and the Nagoya Protocol will meet again in 2020 in Kunming, China, at which time further decisions will be made not only on DSI but also on the CBD's post-2020 global biodiversity framework (Post-2020 Framework). The issue of DSI has been plugged into intersessional activities supporting the development of the Post-2020 Framework, making it central to how DSI will be dealt with in the future. Many developing countries have announced that they will only agree a Post-2020 Framework if it provides solutions with regard to DSI and benefit-sharing.

But what are we actually talking about when we refer to DSI? What is DSI and how is it used? What does this have to do with the CBD and the Nagoya Protocol? And why is this topic so controversial? This Introductory Guide aims to help African policy-makers and stakeholders to understand the discussion around DSI and the positions of various actors.

Part 1 of this paper explains why DSI is an important issue for the CBD and Nagoya Protocol and it identifies the positions of various actors that have developed over the last couple of years.

Part 2 looks at the international context of the discussions on DSI. It traces the steps taken so far by the international community to address the issue and looks at what has been agreed upon. It also fits the issue of DSI into the wider context, including the Post-2020 Framework, the Sustainable Development Goals (SDGs) and other international instruments and fora.

Part 3 is mainly intended for readers who are not familiar with the science behind the discussions on DSI. It delves deeper into the technical side of things, examines what might be understood by the concept and looks at various uses of DSI. **If you are not familiar with these aspects of the discussion, you may wish to read Part 3 first before reading the rest of this Introductory Guide.**

PART 1: DIGITAL SEQUENCE INFORMATION ON GENETIC RESOURCES – WHAT'S THE ISSUE?

Sequencing technology, biological research and biotechnology

The technologies and methods used in science are constantly evolving. This is also true of the biological sciences, which have been heavily influenced in recent years by advances in computer technology and the associated "big data" revolution. Of relevance to the discussion on DSI are:

- faster and cheaper "next generation" and "third-generation" sequencing technologies;
- vastly increased capacities to store, analyse and share data;
- widespread use of DSI for research and development, including by commercial actors.

Sequencing technologies: In the past, unravelling the genetic information contained in living organisms was a slow and expensive process. According to the US National Human Genome Research Institute, sequencing the first human genome (which has around 3 million base pairs) took around 4 years and reportedly cost around 450 million USD for the sequencing work as such. By late 2015, a human genome could be sequenced for less than 1,500 USD. Today, there are companies working to reduce this cost to less than 100 USD. Some scientists do sequencing in their own laboratories and others have it done for them by commercial service providers. Small handheld nanopore sequencers that can be used in the field cost around 1,000 USD for a starter pack, including technical support.

Routine sequencing of deoxyribonucleic acid (DNA), ribonucleic acid (RNA) and proteins (amino acids) has become commonplace for scientists in many fields of biology. These technologies are now also used in everyday applications, such as food quality control, identifying the species of fish used to make sushi, detecting invasive organisms in water bodies, diagnosing plant and animal pests and diseases, verifying the identity of natural products, and many others.

Data storage and data-sharing: An enormous amount of biological data is being generated as a result of sequencing. There are already petabytes (1015 bytes = PB) of data stored in various databases and personal computers around the world. One of the major databases, the European Bioinformatics Institute (EMBL-EBI), for example, can already store over 155 PB of data alone and the influx of data is growing faster than the ability to store and process it.[1]

There are reports that there are as many as 1,700 databases and repositories of biological data and associated information around the world. What types of biological data are we talking about? DNA, RNA and proteins have already been mentioned. EMBL-EBI, for example, offers a range of different information and data on genomics, proteins, gene expression, small molecules, protein structures, metabolic systems, ontologies and related literature. Other databases are specialized in storing certain types of data and information, for example, proteins, pathogens, genome sequences of crop plants, human sequences etc.

Powerful online search tools and widespread internet connectivity mean the data can be located quickly, easily transferred from one place to another and downloaded for analysis. Data are uploaded to and retrieved daily from different databases around the world but also shared among researchers directly. This means that data are often replicated, i.e. stored in multiple different places (and possibly in different forms), including in private databases at companies, publicly accessible databases or journal repositories, etc. Companies from various industry sectors, such as pharmaceuticals, plant breeding etc. have their own proprietary data stored in private databases, although it is not known how much data and information is held by the private sector.

An enormous amount of data is stored in publicly accessible databases, most of which are located in developed countries. Three large databases are considered to be the main actors: EMBL-EBI (mentioned above), the DNA DataBank of Japan (DDBJ) and GenBank, which is hosted by the National Center for Biotechnology Information (NCBI) in the USA. Together, these databases form the International Nucleotide Sequence Database Collaboration (INSDC). They have open access policies and exchange their data daily.

1 Cook et al. 2018. The European Bioinformatics Institute in 2017: data coordination and integration. Nucleic Acids Research 46, D1: D21-D29. https://academic.oup.com/nar/article/46/D1/D21/4658830.

Data use by commercial actors: DSI is used for basic, applied and commercial research for a wide range of purposes. Most of the publicly accessible data and information are used for non-commercial purposes but companies are also known to use these data and the information, for example, to optimise sequences used for product development by searching and comparing thousands of similar sequences and identifying the sequences responsible for desired traits. In some cases, these companies may actively seek intellectual property protection in the form patents on inventions based on such data and information. For example, sequences obtained from databanks can be synthesised and used for the development of patentable vaccines. As DSI is widely used in biological research, there are many examples from basic, applied and commercial research in a number of fields ranging from research on pathogens through to the use of CRISPR-CAS9 in plant breeding and the emerging field of synthetic biology. For more detail on the use of DSI, see Part 3 of this Information Guide and refer to the scoping studies on DSI prepared for the Commission on Genetic Resources for Food and Agriculture[2], the International Treaty on Plant Genetic Resources for Food and Agriculture[3], and the CBD[4].

DSI and the CBD objectives

There are many good things about data-sharing and the availability of publicly accessible biological data, which is one of the reasons why many millions of dollars of public funding are made available to support public databases. Among other things, it promotes collaboration, enables replication of work or the application of data in new contexts, and it uses resources efficiently. Many researchers in both public institutions and private companies are very much in favour of having access to enormous amounts of biological data for research and development.

The availability of data in databases around the world makes it possible for researchers to access and use biological data for their research without needing to access the original biological material or sequence it themselves. To put this into perspective, EMBL-EBI, for example, receives 3.3 million unique site visits to its websites every month and there are over 38 million requests made to the websites each week.

All Parties to the CBD recognize that the use of DSI contributes significantly to fulfilling the first two objectives of the CBD, namely by supporting conservation of biodiversity and the sustainable use of its components, and that DSI is important for research and development, especially in fields such as food security and health[5]. There is also recognition that many developing countries do not have the necessary technical, institutional and human capacity to make full use of the potential offered by the vast amount of publicly available DSI and that this gap might be partly addressed through enhanced capacity building and technology transfer. There is, however, significant debate about the implications of the use of DSI for the third objective of the CBD and the objective of the Nagoya Protocol, namely fair and equitable sharing of benefits arising from the utilisation of genetic resources.

2 Heinemann, J.A., et al.. 2017. Exploratory Fact-finding Scoping Study on "Digital Sequence Information" on Genetic Resources for Food and Agriculture. Background Study Paper 68. Food and Agriculture Organization.

3 Welch, E.W., et al. 2017. Potential implications of new synthetic biology and genomic research trajectories on the International Treaty for Plant Genetic Resources for Food and Agriculture. Food and Agriculture Organization.

4 Laird, S.A. et al. 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, 2018. Convention on Biological Diversity. CBD/SBSTTA/22/INF/3 CBD/DSI/AHTEG/2018/1/3.

5 COP14 Decision 14/20 Digital sequence information on genetic resources. https://www.cbd.int/decisions/cop/?m=cop-14

Diverging positions on access and benefit-sharing for DSI

The developments in science outlined above have created a situation in which publicly accessible biological data are regularly used for research and development, including by commercial actors, without DSI users having to directly access the original biological material, obtain prior informed consent (PIC) from a provider and share benefits arising from utilisation of the samples on mutually agreed terms (MAT).

There is disagreement about whether ABS obligations should apply to the use of DSI. The central elements of the discussion include:

- whether the definition of genetic resources covers DSI or not;
- whether DSI should be regulated;
- whether it would be practical to require PIC for access to data and/or information;
- whether benefit-sharing obligations arise from the use of DSI in situations where PIC is not required and MAT have not been established; and
- whether open access to DSI is a sufficient form of benefit-sharing or can be made sufficient through capacity development and technology transfer.

The definition of "genetic resources"[6] in the CBD has been used in the discussion to argue both for and against the inclusion of DSI in the access and benefit-sharing (ABS) regime. This has mostly focused on the meaning of the word "material", used in the CBD definition of genetic resources, and whether this should be understood to include only physical biological samples or also DSI. A number of Parties to the CBD and its Nagoya Protocol have taken the position that DSI is not covered by the definition and is therefore excluded from the scope of these instruments. Most of these Parties are developed countries, which might be expected to benefit most from the existence of freely accessible biological data. Other actors, including many developing and biodiversity-rich countries, maintain the position that genetic material contains genetic information and that DSI falls within the definition of genetic resources.

Some actors argue that **DSI has become so important for modern biological sciences and biotechnological innovations** that its use should not be regulated because of the potential societal benefits of its use, which would be negatively impacted. Although the Parties to the CBD recognise these societal benefits, some, again mostly developing countries, argue that societal benefits are not a sufficient reason to exclude benefit-sharing obligations for the use of DSI. The African Group[7], for example, uses patent-protected innovations as an example. Such inventions may benefit society but patent owners are still financially rewarded when their inventions are used by others.

The practicality of requiring PIC for access to DSI is another aspect of the discussions that has generated strong reactions to date, especially from users of DSI, such as researchers in the academic and private sectors, as well as administrators of large publicly funded databases and a significant number of governments. They argue that the way in which DSI is produced, stored, transmitted and used make it impossible to control access to data and require PIC. These actors also point out that governmental policies, scientific practice, and requirements by journals, funding bodies and intellectual property authorities make it necessary for them to publish their data.

Many developing countries have acknowledged the potential benefits of open access, the possible difficulties with **regulating access to DSI** and that such regulation could be detrimental to research and development. At the same time, they do not accept that biological data should continue to be used freely in the absence of benefit-sharing obligations or that benefit-sharing obligations would only arise through having access requirements for DSI. These countries maintain the view that benefit-sharing is meant to provide economic and financial resources to help developing countries rich in biodiversity pursue development policies that support sustainable use and conservation of biodiversity, and they are concerned that the use of DSI in the absence of benefit-sharing would undermine the objectives of the CBD and the objective of the Nagoya Protocol.

That open access to DSI constitutes sufficient benefit-sharing is an argument that has repeatedly been advanced by some participants in the discussions on DSI, particularly by the research community. However, actors who lack the capacity to use DSI cannot benefit from it, even if it is, in theory, accessible. This clearly points to the need for capacity development and technology transfer, especially in Africa. So far, no resources have been made available for this purpose through the CBD. Furthermore, it is not clear whether capacity development and technology transfer alone will sufficiently address the "gap" between the capacities of actors in developing and developed countries.

6 "Genetic resources" means genetic material of actual or potential value. "Genetic material" means any material of plant, animal, microbial or other origin containing functional units of heredity (Art. 2 CBD).

7 Submission by the African Union Commission on behalf of the African Group to the SCBD dated 31 May 2019. https://www.cbd.int/abs/DSI-views/2019/AfricanGroup-DSI.pdf

Positions and interests of various DSI stakeholder groups

The stakeholders who have an interest in the outcome of the international discussions on DSI include:

- those responsible for making decisions at the international level, including Parties to the CBD and the Nagoya Protocol;
- other governments, notably the USA;
- those with an interest in the potential impacts on benefit-sharing;
- policymakers and regulators responsible for implementing ABS systems at the national level;
- the research community, as both producers and users of DSI, as well as research associations and funding bodies;
- industry actors, as both producers and users of DSI, including but not limited to sectors such as biotechnology, animal breeding, plant breeding, cosmetics, biocontrol and biostimulants, food and feed, and pharmaceuticals, as well as industry or business associations representing these actors;
- database operators, who are central actors with respect to the storage, processing and transmission of large amounts of DSI;
- civil society organizations; and
- indigenous peoples and local communities.

A key task for all involved will be to thoroughly understand the real interests of the abovementioned stakeholder groups and to identify solutions to the issues that can adequately address their concerns, which could be challenging given the diverging views on the matter. Any solutions need to be developed in both the context of the discussion on DSI and the process of developing the Post-2020 Framework.

During the 2017-2018 and 2019-2020 intersessional periods between the COP and COP-MOP, a number of Parties to the CBD and other stakeholders have made submissions on DSI. The wide range of positions and arguments of these actors are reflected in these submissions, which can be viewed on the CBD Secretariat's website.[8]

Approaches to dealing with DSI

Many developing countries have clearly indicated that they are not prepared to accept a situation in which the use of DSI continues without benefit-sharing obligations. Because there is no international agreement on how DSI should be handled, various national approaches are emerging as countries attempt to address DSI through the bilateral ABS system. This includes, for example, including references to DSI in national legislation, regulations and policy, and/or inclusion of DSI in bioprospecting permits, prior informed consent (PIC) applications, mutually agreed terms (MAT), and material transfer agreements (MTA). Emerging national measures include the need to obtain permission to generate sequences from samples of biological material, restrictions on the distribution, use and publication of the resulting data, and benefit-sharing obligations. Some countries have indicated that continued opposition to benefit-sharing could result in more restrictions on the utilization of physical biological samples.

Nevertheless, there will be many challenges associated with regulating DSI through bilateral approaches, including monitoring and compliance associated with data use around the world. Some Parties, including the African Group[9], have suggested that a multilateral approach to dealing with DSI would make more sense, e.g. adopting a global benefit-sharing mechanism under Article 10 of the Nagoya Protocol. Article 10 allows parties the opportunity to think about the need for and potential "modalities" of a multilateral mechanism to address benefit-sharing in certain situations, i.e. when genetic resources and the associated traditional knowledge occur in transboundary situations or it is not possible to grant or obtain PIC for the utilisation of these resources.

8 https://www.cbd.int/abs/dsi-gr.shtml

9 Submission of Ethiopia on behalf of the African Group to the SCBD dated 8 September 2017 https://www.cbd.int/abs/DSI-views/Ethiopia-AU-DSI.pdf

PART 2: THE INTERNATIONAL PROCESS

DSI in the CBD forum

In 2016 in Mexico, at the thirteenth CBD Conference of the Parties (COP13), DSI issue was moved out of the discussions on synthetic biology and into a dedicated work stream. It was agreed that the process would be dealt with in the CBD forum and that this process would also serve the Nagoya Protocol. Two parallel decisions[10] were made that put an information-gathering process in place, involving a scoping study on DSI and its use as well as the submissions of views and information by interested actors on the potential implications of the use of DSI for the objectives of the CBD and the Nagoya Protocol. This information was reviewed by an Ad-Hoc Technical Expert Group (AHTEG), which met in early 2018. The outcomes of this AHTEG were considered at the Subsidiary Body on Scientific, Technical and Technological Advice (SBSTTA) in July 2018 and a recommendation was developed for consideration by CBD COP (COP14) and the third Conference of the Parties serving as the Meeting of the Parties to the Nagoya Protocol (COP-MOP3) in Sharm El Sheik, Egypt in November 2018. At COP14, the following key decisions were made, which are relevant to the discussions on DSI:

- Decision14/20[11] – Digital sequence information on genetic resources; and
- Decision 14/34[12] – Comprehensive and participatory process for the preparation of the post-2020 global biodiversity framework.

Both of these decisions were acknowledged and welcomed by the Parties to the Nagoya Protocol in decision COP-MOP Decision 3/12.[13]

DSI and the science-policy process – the intersessional period 2018-2019

In Egypt, the Parties decided to put in place a science-policy process to inform further discussions on DSI. Decision 14/20 acknowledges:

- the different positions on DSI;
- the relevance of DSI for all three objectives of the CBD, scientific research and other non-commercial and commercial activities;
- the differences in the capacity of different actors around the world to access, use, generate and analyse DSI. The decision also encouraged capacity-building and technology transfer to address this gap;
- the growth in the generation and use of DSI, its publication in both public and private databases, advances in data analytics and that new technologies play an important role for current and future utilization of genetic resources;
- the challenges associated with linking DSI to the biological sample from which it was derived and the continuously evolving media in which information is stored and shared; and
- that some countries have already adopted domestic measures that regulate access to and use of DSI as part of their ABS frameworks.

The agreed process to inform discussions at COP15 in Kunming in 2020 includes the following elements:

- the submission of views and information by Parties and other stakeholders on the concept of DSI (including the term DSI and its scope), whether ABS regimes at the domestic level deal with the use of DSI or benefit-sharing arrangements arising from its commercial and non-commercial use, and capacity-building needs regarding the access, use, generation and analysis of DSI;
- establishment of an extended Ad Hoc Technical Expert Group; and
- preparation of four peer-reviewed fact-finding studies, which will be finalized by the end of 2019, on the concept and scope of DSI, traceability of digital information, public and private databases holding DSI, and domestic measures on use of DSI for R&D and benefit-sharing arising from its use.

10 COP13 Decision XIII/16 https://www.cbd.int/doc/decisions/cop-13/cop-13-dec-16-en.pdf and COP-MOP2 Decision 2/14 https://www.cbd.int/doc/decisions/np-mop-02/np-mop-02-dec-14-en.pdf
11 https://www.cbd.int/doc/decisions/cop-14/cop-14-dec-20-en.pdf
12 https://www.cbd.int/decision/cop/?m=cop-14
13 https://www.cbd.int/doc/decisions/np-mop-03/np-mop-03-dec-12-en.pdf

In early 2020, the extended Ad Hoc Technical Expert Group will meet to consider the views and information and the peer-reviewed studies, develop options for operational terms and their implications to provide conceptual clarity on DSI, especially on the scope of DSI and its use, and identify key areas for capacity-building.

The outcomes of this meeting will be submitted to the Open-Ended Working Group (OEWG) that was established to consider the Post-2020 Framework. This working group will make recommendations to COP15 on how to address DSI in the context of the Framework.

The Post-2020 Global Biodiversity Framework

In 2010, the Strategic Plan for Biodiversity 2011-2020 and its Aichi targets were adopted by the tenth meeting of the COP in Nagoya, Japan. The Strategic Plan provides an overall framework on biodiversity for the entire UN system.

Parties translated the Strategic Plan and the Aichi targets into national biodiversity strategies and action plans, but most of these goals and targets have not yet been met. The Parties decided at COP14[14] to establish a comprehensive and participatory process for the preparation of a Post-2020 Framework that will provide guidance on long-term strategic directions to achieving the 2050 Vision for Biodiversity, i.e. living in harmony with nature. It is envisaged that the Post-2020 Framework will be adopted by the CBD COP and given endorsement by the COP-MOPs of the Nagoya Protocol and Cartagena Protocol.

The new framework should seek to identify and strengthen synergies with other international instruments and explore different options for gaining high-level attention for biodiversity-related issues, e.g. how biodiversity links to fulfilment of the Sustainable Development Goals (SDGs). The UN General Assembly has decided to convene a summit on biodiversity before COP15 to highlight the link between biodiversity and the 2030 Agenda for Sustainable Development, as well as the urgency of action with respect to biodiversity loss and the need for high-level support for the Post-2020 Framework.

The preparation of the Post-2020 Framework is intended to be a transparent, open and step-by-step process built on consensus. It encourages a wide range of different stakeholders, including IPLCs, to get involved and contribute to the process, which is guided by principles such as inclusiveness, transparency, gender responsiveness etc. The process is intended to be knowledge-based and a wide range of possible sources of information and inputs have been identified. Stakeholders can contribute to the process through consultations, online discussions, submissions etc.

14 COP14 Decision 14/34 https://www.cbd.int/decisions/cop/?m=cop-14

Intersessional and subsidiary bodies of the CBD will also address particular aspects of the Post-2020 Framework, including the Working Group on Article 8j of the CBD, SBSTTA etc. The process is also being supported by the OEWG, which will meet three times in the run up to COP15.

Importantly, the OEWG will consider the outcomes of the science-policy process on DSI and make recommendations to the COP on how to address DSI in the context of the Post-2020 Framework.

ABS and the Post-2020 Global Biodiversity Framework

Aichi Target 16 states that by 2015, the Nagoya Protocol is in force and operational, consistent with national legislation. The Nagoya Protocol came into force in 2014 and as of June 2019, it had 117 Parties. This means that Aichi Target 16 has partially been achieved. However, many Parties are still in the process of establishing their ABS frameworks and institutional structures to enable the implementation the Protocol at the national level. In other words, the full operationalization of the Nagoya Protocol has not yet been achieved.

At the COP-MOP3[15] in Egypt, the Parties to the Nagoya Protocol welcomed Decision 14/34 on the Post-2020 Framework. COP-MOP Decision 3/15[16]:

- encourages Parties to the Nagoya Protocol to undertake measures to enhance the implementation of the Nagoya Protocol in the context of the Post-2020 Framework;
- invites Parties to the Nagoya Protocol to participate in the process leading up to adoption of the Post-2020 Framework; and
- recommends that findings on general issues of compliance as well as the outcomes of the first assessment and review of the effectiveness of the Protocol be considered in the development of the Post-2020 Framework. The Compliance Committee was also asked by the COP-MOP to give more thought to how to promote compliance with the Nagoya Protocol in the Post-2020 Framework.

15 COP-MOP3 Decision NP-3/12. https://www.cbd.int/doc/decisions/np-mop-03/np-mop-03-dec-12-en.pdf
16 COP-MOP Decision NP-3/15. https://www.cbd.int/doc/decisions/np-mop-03/np-mop-03-dec-15-en.pdf

ABS, DSI and the Sustainable Development Goals

In September 2015, the United Nations adopted the "2030 Agenda for Sustainable Development", with its 17 SDGs and 169 targets to be reached by 2030. The Agenda addresses poverty alleviation, equity, social justice and sound environmental management to ensure economically, ecologically and socially sustainable development. Unlike the Millennium Development Goals, it sets a global agenda for sustainable development, meaning that the focus is not only on developing countries.

Biodiversity and functioning ecosystems are fundamental to fulfilling many of the SDGs. ABS also has a role in the fulfilment of the SDGs, both directly and indirectly through:

- support for health-related research;
- encouraging the development of biodiversity-based value chains and economic development in local communities;
- fostering innovative use of genetic resources and the associated traditional knowledge;
- transfer of technology;
- development of effective, accountable and transparent institutions, including in R&D chains;
- conservation and sustainable use of wild genetic resources on land and from the ocean;
- conservation and sustainable use of genetic diversity of cultivated plants and domesticated animals, thus contributing to food security, improved nutrition and climate change adaptation; and
- strengthening international partnerships through research.

Discussions around DSI in the CBD forum have highlighted that use of DSI is fundamental to research on biodiversity conservation, sustainable use, food security, health etc. In this sense, DSI can also be seen as supporting achievement of a wide range of the SDGs, including both biodiversity related and non-biodiversity related goals.

This indicates that ABS and DSI are both relevant for the achievement of the SDGs. However, if the use of DSI potentially undermines ABS, there is also a potential for negative impacts not only on the objectives of the CBD and the objective of the Nagoya Protocol but also on fulfilment of the SDGs. This potential contradiction should be considered and addressed by the Post-2020 Framework.

DSI in other fora

The discussions around DSI are not restricted to the CBD and its Nagoya Protocol. Other international fora are discussing DSI and trying to understand the potential implications of the use of DSI for the relevant international instruments. These include:

- the Food and Agriculture Organization of the United Nations (FAO), its Commission on Genetic Resources for Food and Agriculture, and the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA);
- the World Health Organization (WHO), its Pandemic Influenza Preparedness Framework (PIP Framework) and related discussions around ABS for human pathogens;
- the United Nations Convention on the Law of the Sea (UNCLOS) and the current discussions under the auspices of the UN General Assembly on a legally binding international instrument on biodiversity beyond national jurisdiction; and
- the World Intellectual Property Organization (WIPO) and its Intergovernmental Committee on Genetic Resources, Traditional Knowledge and Folklore.

The PIP Framework, which was adopted in 2011, makes explicit reference to "genetic sequence data" and the need to consider the implications of its use for the benefit-sharing system it establishes.

The issue of DSI first came onto the international agenda of the ITGRFA in 2013 in the context of discussions about enhancing the Multilateral System. In November 2019, the eighth meeting of the Governing Body of the Treaty (GB8) will consider the outcomes of nine formal meetings of a working group on this topic and numerous informal gatherings, which have taken place over the past six years. DSI is the main outstanding issue preventing consensus on the adoption of a new Standard Material Transfer Agreement (SMTA), which has been one of the central themes in these negotiations. The current position of the African Region is that it will not accept a revised SMTA that does not contain provisions on ABS for DSI but would be willing to accept an instrument and/or GB8 decision referring to "associated information, including genetic sequence data" and not specifically DSI.

These fora are not working in isolation but are looking to see how others are trying to deal with the issues, especially in the CBD. This is important for ensuring consistency in the way the international community deals with DSI. To an extent it also prevents duplication of work, although each instrument eventually needs to find solutions for its particular area of focus.

Looking ahead to Kunming, 2030 and beyond

Although the focus in the CBD forum is on the implications of the use of DSI for the objectives of the Convention and its Nagoya Protocol, it is important to keep the links to other instruments mentioned above in mind as it highlights the cross-cutting nature of DSI and the need for solutions that are consistent across the entire UN system. The Post-2020 Framework will play an important role in this regard.

The process leading up to adoption of the Post 2020-Framework should clarify how the international community will deal with the issue of DSI. Key questions that need to be asked by actors on the way to COP15 and in the development of the Post 2020 Framework include:

- How will the Post-2020 Framework address the three objectives of the Convention in a balanced way?
- What role should ABS play in the Post-2020 Framework, including questions about strategic goals and targets on implementation, compliance etc.?
- What is the link between ABS, DSI and the SDGs?
- How will DSI, including benefit-sharing, be addressed?

Many developing countries have indicated that they **will not support the adoption of a Post-2020 Framework that does not contain a benefit-sharing solution for DSI.**

PART 3: WHAT IS DSI AND HOW IS IT USED?

What is DSI?

In order to understand the arguments of the different actors with respect to DSI, it is important to understand what we are talking about when we refer to DSI.

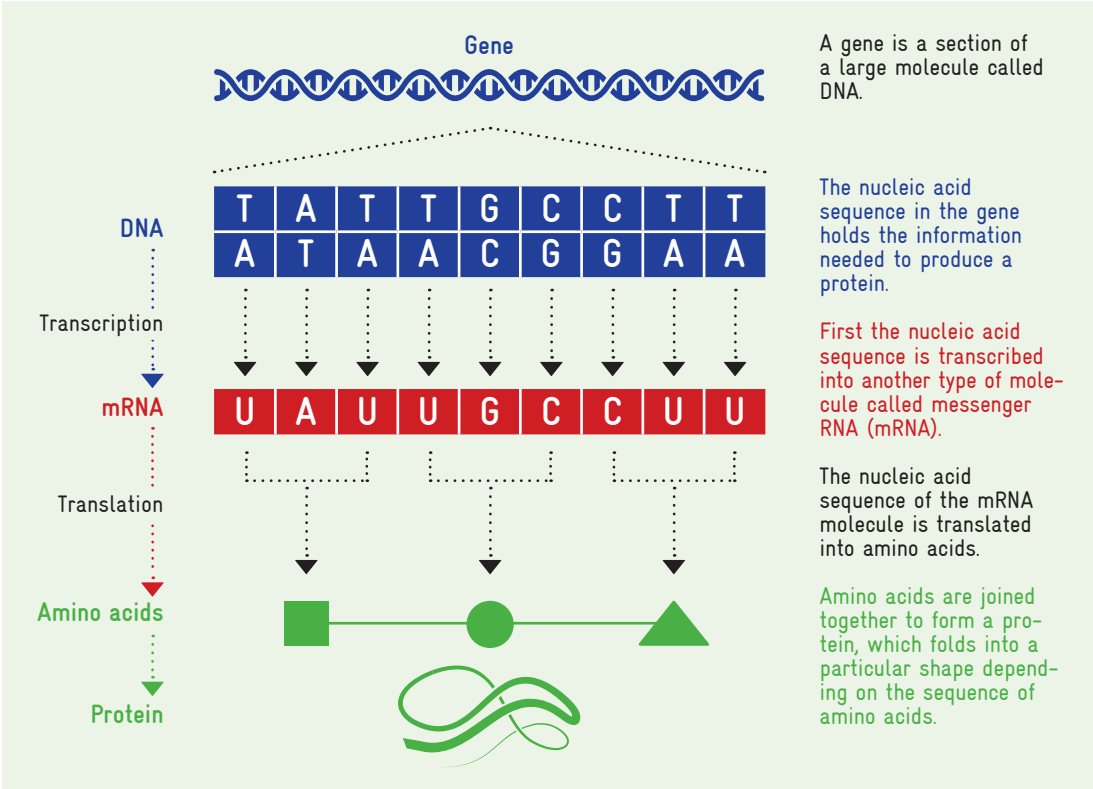
What are we actually talking about? This is a good question and unfortunately, there is no exact answer to this question yet as no accepted definition of DSI has been developed. The acronym DSI has found its way into the international negotiations of the CBD and is being used as a placeholder until such time as an alternative is agreed upon. It needs to be kept in mind that there is no universal understanding of its meaning and scope and that actors could possibly be referring to (or understanding) different things.

Having said this, it is still possible to explore the concept of DSI and to look into what the term DSI might cover. This is important to understand as the scope of the concept will determine the scope of the political discussions and any decisions made in relation to DSI.

Which biochemical molecules are relevant to the discussion on DSI?

So far the international discussions on DSI seem to focus on macromolecules built from a number of distinct building blocks. Such sequential molecules are, for example, DNA, RNA and proteins. In the image below, you can see the process by which the information in the genome is used for the production of proteins. There are also other sequential macromolecules, such as polysaccharides. Sequential information could also describe the sequence of atoms in a biochemical compound expressed by a genetic resource.

Human genetic resources were explicitly excluded from the scope of the CBD by a decision taken at COP 2, so we can be confident that the discussions on DSI do not extend to human sequences.



How is DSI generated?

In order to generate DSI, a small sample of biological material is needed and the process below would be followed.

Using a protocol, the molecule of interest is isolated, e.g. DNA, RNA or the proteins.

↓

Chemical and physical methods are used to determine the building blocks of the molecules and their order.

↓

The raw data is cleaned and the data analysis is conducted using other sequences for comparison. Analysis involves the use of computer algorithms and scientific expertise.

↓

The results of the analysis and the associated data may be published, e.g. in a database or journal repository.

DSI could include datasets providing different levels of information, starting with raw sequences resulting from the sequencing process, sequences that have been "cleaned", i.e. all errors have been removed, sequences that have been analysed and possibly annotated using algorithms, and sequences that have been analysed by a researcher and the function has possibly been identified.[17]

The outcome of this process might look like this published partial mRNA sequence of the Conus snail (extract from GenBank):

Conus regius conotoxin Rg3.18 precursor, mRNA, partial cds

GenBank: MF588952.1

GenBank Graphics

```
>MF588952.1 Conus regius conotoxin Rg3.18 precursor, mRNA, partial cds
ATGATGAGCAAACTGCGCGTGCTGCTGACCATTTGCCTGCTGCTGTTTCCGCTGAGCGCGCTGCCGCTGG
ATGGCGATCAGCCGGCGGATCAGCCGGCGAAACGCATGTGGAACGGCAAACTGGCGGCGCGCAAACCGCG
CTTTGATAAATATGATCTGGTGCGCGGCTGCTGCCCGCCGCAGTGGTGCGGCCCGGATTGCACCAGCCCG
TGCTGCGGC
```

17 Jaspars, M. in Laird, S.A. et al. 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, 2018. Convention on Biological Diversity. CBD/SBSTTA/22/INF/3 CBD/DSI/AHTEG/2018/1/3.

DSI in databases is structured using particular formats and conventions. In addition to the sequences themselves, the databases include other information, such as metadata about the organism (origin, date ecological relationships, environmental factors etc.), the sequencing method used, the relevant publication etc. The example below is for a partial mRNA sequence of the Conus snail (extract from GenBank):

conotoxin Rg3.18 precursor, partial [Conus regius]

GenBank: AUJ88076.1

Identical Proteins FASTA Graphics

Go to:

```
LOCUS       AUJ88076                  73 aa            linear   INV 08-JAN-2018
DEFINITION  conotoxin Rg3.18 precursor, partial [Conus regius].
ACCESSION   AUJ88076
VERSION     AUJ88076.1
DBSOURCE    accession MF588952.1
KEYWORDS    .
SOURCE      Conus regius
  ORGANISM  Conus regius
            Eukaryota; Metazoa; Lophotrochozoa; Mollusca; Gastropoda;
            Caenogastropoda; Hypsogastropoda; Neogastropoda; Conoidea; Conidae;
            Conus; Stephanoconus.
REFERENCE   1  (residues 1 to 73)
  AUTHORS   Franco,A., Dovell,S., Moller,C., Grandal,M., Clark,E. and Mari,F.
  TITLE     Structural Plasticity of Mini-M Conotoxins: Expression of all
            mini-M subtypes by Conus regius
  JOURNAL   FEBS J. (2017) In press
   PUBMED   29283511
  REMARK    Publication Status: Available-Online prior to print
REFERENCE   2  (residues 1 to 73)
  AUTHORS   Mari,F. and Grandal,M.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2017) Marine Biochemical Sciences, NIST, 331 Fort
            Johnson Rd, Charleston, SC 29412, USA
COMMENT     ##Assembly-Data-START##
            Assembly Method       :: Trinity v. 2/2017
            Sequencing Technology :: Illumina
            ##Assembly-Data-END##
FEATURES             Location/Qualifiers
     source          1..73
                     /organism="Conus regius"
                     /db_xref="taxon:101314"
     Protein         1..>73
                     /product="conotoxin Rg3.18 precursor"
     Region          3..68
                     /region_name="Conotoxin"
                     /note="pfam02950"
                     /db_xref="CDD:308546"
     CDS             1..73
                     /coded_by="MF588952.1:1..>219"
ORIGIN
        1 mmsklrvllt iclllfplsa lpldgdqpad qpakrmwngk laarkprfdk ydlvrgccpp
       61 qwcgpdctsp ccg
//
```

Examples of use of DSI

It is important to distinguish between different types of uses of DSI. Different types of sequences are used in a wide range of scientific disciplines, ranging from basic research through to applied and commercial research. DSI is regularly used as a tool, which provides information that would not be regarded as research and development. Examples include the identification of known microbes, including the diagnosis of pathogens in animals, plants and people or screening foods for safety, identification of genetically modified organisms, sex identification, label certification and/or verification of label ingredients in food products.

DSI is also used widely for basic research, e.g. in taxonomic work such as investigating the relationships within and between species or taxa, to validate previous research and to generate new taxonomic knowledge. DNA barcodes, for example, are widely used. DNA barcodes are short sections of DNA from a specific gene or genes, which can be compared with reference sequences to identify species. There are various examples in the literature of DNA barcodes being used for research that supports conservation or sustainable use of biodiversity.

Example: Identifying shark species in meat products

Sharks are caught and sold in a variety of commercial products. In one study, researchers[18] tested thirty-five shark meat products, including fillets, jerky, soup, and cartilage pills using DNA barcodes. These sequences were compared with sequences in databases. Not all products could be identified to the species level but the tests nonetheless revealed that many of the products were mislabelled and several contained species listed in the Appendices of the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES). Even more of the products included species considered to be near-threatened, vulnerable or endangered species according to the International Union for the Conservation of Nature (IUCN) Red List.

18 Hellberg, R.S., Isaacs, R.B. and Hernandez, E.L. 2019. Identification of shark species in commercial products using DNA barcoding. Fisheries Research. 210: 81-88.

DSI is also used for applied sciences, such as plant and animal breeding to predict breeding value of individual animals. It is also possible, for example, to use DSI to modify genes, which are subsequently inserted into bacteria so that they produce desirable high-value molecules. There are many examples of transgenic organisms that have been developed to produce ingredients for medicines or other industrial ingredients. This can replace farm-based production of traditional medicinal plants through industrial fermentation or by farming transgenic plants.

Example: DNA sequences and genetic engineering of plants

In one study, Researchers[19] used four published DNA sequences derived from the plant *Artemisia annua* to engineer a metabolic pathway in a genetically modified tobacco plant. The purpose of this was to create tobacco plants that produce artemisinic acid, which is a precursor of the anti-malarial drug Artemisinin. One gene sequence, FPS, was published by Chinese researchers from *Artemisia annua* plants collected from Sichuan Province, China[20]. The sequence of the ADS-gene was published by public and private researchers from Sweden and the Netherlands[21]. The authors of this research did not indicate the source of the genetic resource but the nucleotide sequence data was deposited in the GenBank/EMBL Database. The sequences of the CYP and the CPR genes were published by a team of researchers from public institutions and companies in the USA[22]. The sequences were derived from seeds bought from a UK company which "collects, processes and sells seeds of trees and shrubs, seeds of Mediterranean, subtropical and tropical plants, seeds of ornamental grasses and perennials, seeds of herbs, medicinal and aromatic plants". These DNA sequences were uploaded to GenBank.

19 Fuentes, P. et al. 2016. A new synthetic biology approach allows transfer of an entire metabolic pathway from a medicinal plant to a biomass crop. eLife 5. DOI: 10.7554/eLife.13664

20 Chen D-H et al. 2000. Expression of a chimeric farnesyl diphosphate synthase gene in Artemisia annua L. transgenic plants via Agrobacterium tumefaciens-mediated transformation. Plant Science 155: 179–85.

21 Mercke P et al. 2000. Molecular cloning, expression, and characterization of amorpha-4,11-diene synthase, a key enzyme of artemisinin biosynthesis in Artemisia annua L. Archives of Biochemistry and Biophysics 381 (2): 173-80.

22 Ro D-K et al. 2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature 440(13): 940-3.

There has also been rapid growth in "cut and paste" gene editing techniques, for example with CRISPR-Cas9, which are supported by DSI. These techniques allow researchers to edit genes with unprecedented ease and accuracy. Although these technologies are still under development and primarily used in the medical sector, they have the potential to transform many other fields of the biological sciences. In plant breeding, for example, it is expected that gene editing plants to make them more resistant to diseases or to produce new metabolites can greatly accelerate improvements that currently rely on slow and incremental conventional breeding techniques.

There are twenty natural amino acids, which can be grouped together to form long chains, i.e. proteins. Proteins are responsible for the dynamic process such as maintenance, replication, reproduction etc. Sequencing technology plays an important role in research looking into the composition, structure, dynamics and function of proteins[23] and supports a wide range of research, including medicinal research.

23 Breda, A. Fonseca Valadares, N., de Souza, O.N. and Garratt R.C. 2008. Protein Structure, Modelling and Applications in Bioinformatics in Gruber A, Durham AM, Huynh C, et al. (eds). Tropical Disease Research: A Practical and Case Study Approach. National Center for Biotechnology Information (US). https://www.ncbi.nlm.nih.gov/books/NBK6824/

Example: Chemical synthesis of proteins from sequences for the development of painkillers

Since 2004, a major pain killer has been marketed as Prialt in the USA, which was developed from ∂-conotoxin proteins isolated from a marine snail collected in the Philippines many years beforehand. Fuentes et al. (2016) obtained protein sequences from the marine snail *Conus regius* in order to synthesize ∂-conotoxins proteins in the laboratory for testing in mice.[24] The purpose of their research was to investigate the prevention of chronic cancer chemotherapy-induced neuropathic pain. The sequences used for synthesis of the proteins had been determined by a team of researchers from public research institutions in the USA and Argentina in 2006 and were obtained from GenBank.

24 Romero, H.K. et al. 2017. Inhibition of ∂9∂10 nicotinic acetylcholine receptors prevents chemotherapy-induced neuropathic pain. Proceedings of the National Academy of Sciences USA

DSI is also used in commercial research. Companies use large amounts of DSI, including DSI from public databases. Some companies conduct basic research, but the public data and information are also used for commercial purposes. Sequences may not, for example, be commercialized directly but are used to support the development of a commercial product, e.g. through optimization of sequences already held by the company.

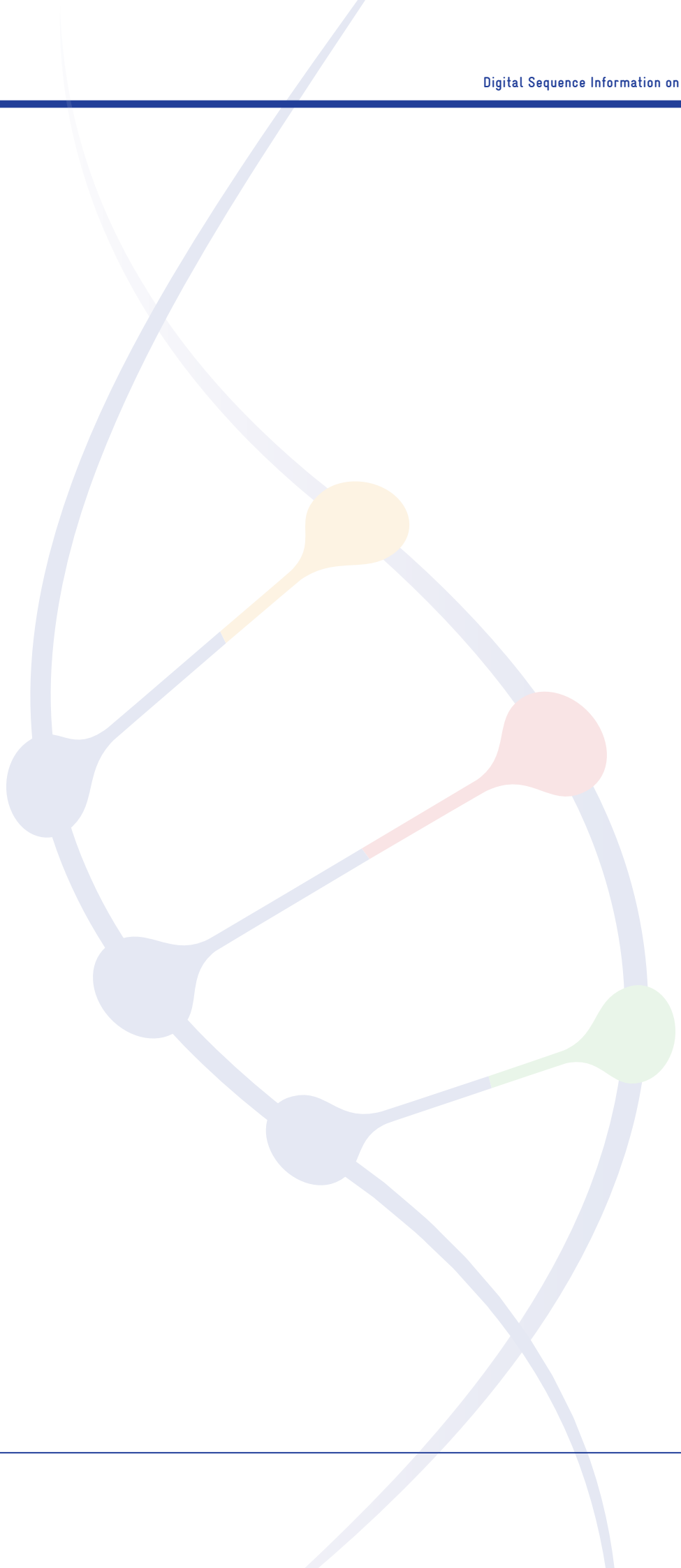
This involves obtaining a large number of similar sequences from a database and using them for comparative purposes.[25] Other companies may seek out and use specific sequences obtained from publicly accessible databanks for product development directly, meaning that sequences could be used to develop patent-protected inventions, such as vaccines. It is, for example, possible to synthesize viruses based on information available in public databases use these in the design and manufacture of medicines or vaccines, rather than obtaining physical samples of viruses.

25 Submission by International Chamber of Commerce to the SCBD dated 15.09.2017. https://www.cbd.int/abs/DSI-views/ICC-DSI.pdf

Example: Use of Ebola genome sequences to develop a drug

Various sequences from the Ebola virus, including full and partial genomes, are available for download from GenBank. Some of these sequences were obtained from blood samples from victims of the 2014 Ebola epidemic in West Africa. In one case, the strain named C15 was isolated from a clinical sample from Guinea by researchers from the Pasteur Institute in Lyon, France, in cooperation with the Nocht Institute in Hamburg, Germany. The Nocht Institute sequenced the C15 strain and uploaded the sequence to GenBank. A US company, Regeneron Pharmaceuticals, subsequently used the sequence to develop a drug against Ebola. Regeneron has since obtained patents on the drug, REGN- EB3, in the US, Nigeria and South Africa, and it has further applications pending in over 100 other countries.[26]

26 Hammond, E. 2019. Ebola: Company avoids benefit-sharing obligation by using sequences. TWN Briefing Paper 99. http://www.twn.my/title2/briefing_papers/No99.pdf



Published by:
Deutsche Gesellschaft für
Internationale Zusammenarbeit (GIZ) GmbH
Bonn and Eschborn, Germany
Dag-Hammarskjöld-Weg 1-5
65760 Eschborn, Germany

T +49 61 96 79-3285
F +49 61 96 79-803285

abs-initiative@giz.de

www.giz.de

Photo: Adobe Stock (©flashmovie; ©sergunt; ©ktsdesign)

Layout/Design: MediaCompany, Agentur für Kommunikation

As at August 2019

GIZ is responsible for the content.