DFG Statement on the Discussion on Digital Sequence Information (DSI)

In October 2014, the *Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization to the Convention on Biological Diversity* entered into force. This internationally legally binding contract, which governs the effective implementation of the Convention on Biological Diversity's (CBD) third objective, is very relevant for international research, as research on biological material is declared as utilization of genetic resources in the CBD as well as in the Nagoya Protocol.

Due to the generalized definition of genetic resources[1] and their utilization[2], almost all areas of Life Sciences whose work requires the exchange of genetic resources across state boundaries are concerned. To emphasize the scope of this concern, some possibly impacted research areas are: Ecological Research, for which animals and plants are collected; Earth Sciences Research, for which soil and water samples are required; Medical Research, for which e.g. blood samples of Zika-Virus-infected humans or special bacterial strains are needed; Fundamental Biological Research, in which model organisms like *Drosophila* or mouse lines often are provided by commercial suppliers or exchanged within the framework of international cooperation.

It goes without saying that a fair and equitable sharing of benefits in basic research has to be self-evident and will be comprehensively supported. The possibilities of a fair benefit sharing in fundamental research stipulated in annex 1 of the Nagoya Protocol (for instance, cooperation, capacity building, etc.) have already been put into practice in respective research projects.

The concrete design of the permission processes, however, is a challenge for researchers. The regulation predefines that the access to genetic resources from provider countries is subject to their permission[3], and further requirements and obligations can be attached to such permissions. Indeed, obtaining the corresponding documents from the provider countries often represents a significant problem for researchers. According to reports from researchers, one of the main issues is identifying the authorities in charge, as these often do not yet exist. In addition, denominating the genetic resources (e.g. with complex soil and water samples) and describing the exact utilization, which frequently does not become evident prior to starting the project, is an issue. Altogether, these

[1] A genetic resource is genetic material of real or potential value, this includes all samples from in-situ sources, including populations of free-living as well as domesticated species and from ex-situ sources (collections of all kinds, public or private)

[2] Utilization is any kind of research and/or developmental activity on the genetic and/or biochemical composition of genetic resources

[3] These encompass a Prior Informed Consent (PIC, an access permission) and Mutually Agreed Terms (MAT, a license agreement)

requirements often lead to massive delays of research projects - even with German Research Consortia with well-established contacts with their partners abroad. These delays can easily encompass periods that make research projects become obsolete. Further, early-career researchers who aim to set up a new project frequently encounter unsolvable problems. The possibility of facilitated access to genetic resources to support research, to which the provider countries are called upon in article 8a of the Nagoya Protocol, offers a constructive solution and should be strongly supported.

Within the scientific community, the described difficulties lead to substantial insecurity with regard to the legal regulations, such that a lot of current research from Germany is preferably carried out in countries where administration is operated smoothly, whereas obtaining genetic resources from other countries is often pursued no further.

While the CBD and the Nagoya Protocol *expressis verbis* refer to material in the sense of genetic resources, the discussion on the equivalency of **digital sequence information (DSI)**[4] was recently opened during the 13th meeting of the Conference of the Parties to the Convention on Biological Diversity and the third meeting of the Parties to the Nagoya Protocol (Cancun, Mexico, December 2016). It was specifically demanded to treat the utilization of digital sequence information as equivalent to the utilization of genetic resources, and that it consequently should be subject to the CBD and Nagoya Protocol regulations. This requirement is a new development, as prior discussions concerning the Nagoya Protocol intentionally did not include digital data.

A development equivalent to this requirement would lead to substantial difficulties for research in the Life Sciences. The foundation of research is based on the publication of new findings and public access to this new knowledge and the underlying data. Research studies that are financed by public funds have to publish their results. Depositing these data in publicly accessible data repositories, the nucleotide databases, is a prerequisite for publishing results based on sequence data analyses. This step is controlled by scientific journals, which do not accept manuscripts for publication without the corresponding reference to data in a database. This mechanism aims at ensuring the replicability and control of research results. Moreover, it creates a tremendous added value, as the freely accessible sequence data represents an essential source of insight for basic research, which allows a sustainable utilization of the generated data and initiates the development and enhancement of new research approaches.

The nucleotide databases are connected within the umbrella organization of all nucleotide databases (International Nucleotide Sequence Database, INSD). This umbrella organization positions itself

[4] However, it has to be assumed that, in the future, also other digital data could be considered as genetic resources.

distinctly for the public accessibility to all accepted data in its *Collaboration Policy* – a fundamental prerequisite that has been accepted by all data providers since the establishment of databases in the 1980s, and which comprises more than 100,000,000 world-wide freely accessible pieces of sequence information to date: "The INSD has a uniform policy of free and unrestricted access to all of the data records their databases contain. Scientists worldwide can access these records to plan experiments or publish any analysis or critique[5]." This general data accessibility, which was elicited by the above-mentioned pressure from scientific journals, represents an enormous resource in basic research and is an important contribution to ensure good scientific practice.

If provider countries set conditions concerning the access to genetic resources to include digital sequence information, these conditions have to be observed and relayed when sharing digital sequence information with a database. At present, the nucleotide databases do not offer any options in this respect, and whether such a development would be considered by database operators is doubtful.

Hence, it can be anticipated that developments leading to restrictions on the availability and accessibility of data and additional administrative requirements will considerably hinder the publication of new findings, and ultimately the funding of new projects. The corresponding research would consequently be affected as well. This again would lead to an irrational development regarding the main mission of the CBD. Basic Research on the understanding of biodiversity, which is essential for its protection, would be substantially impaired by this. Especially research in the provider countries would stand to lose potentially important international collaborations due to the legal insecurities. Researchers in recipient countries will, in the long-run, most probably realign and turn to other research topics and issues that are not afflicted by such administrative constraints.

Provider countries have the right to set conditions for the handling of their genetic resources and their derivatives. If a provider country should demand restrictions concerning the publication of digital sequence information, this can be regulated individually within a license agreement for the project. In the case of such individual restrictions prior to utilization, partners can consider whether a project is to be carried out under these circumstances or not. However, if digital sequence information is generally included into the regulations of the Nagoya Protocol, a general restriction of data utilization would result from this. From the DFG's point of view, data that are to be published must be freely

[5] http://www.insdc.org/policy
International Nucleotide Sequence Database Collaboration Policy
Soren Brunak, Antoine Danchin, Masahira Hattori, Haruki Nakamura, Kazuo Shinozaki, Tara Matise, Daphne Preuss (2002) Nucleotide Sequence Database Policies
Science 298 (5597): 1333 15 Nov 2002

accessible so that research according to established standards will be possible at the international level in the future.

DFG's position on the handling of digital sequence information in brief:

- Project results from research financed by public funds must be published. If digital sequence information are subject to restrictions concerning utilization, publications are not possible. Consequently, restrictions on the utilization of digital sequence information would result in a hindrance of research on respective research topics.

- The legal insecurities concomitant with the described discussions lead to restrictions on the internationality of research. Relevant research demanded in the sense of CBD would be hindered.

- Previously published digital sequence information is part of the public domain. It was published under valid regulations and is unrestrictedly available for future use. For basic research, the published digital sequence information represents an important source of insight in nearly all areas of the Life Sciences and enables new and promising research approaches.