

Digital sequence information–legal questions for patent, copyright, trade secret protection and sharing of genomic sequencing data

C Seitz

Visiting Professor of Law, Faculty of Law and Criminology, University of Ghent, Ghent, Belgium, Lecturer at the Center for Life Sciences Law (CLSL), Faculty of Law, University of Basel, Basel, Switzerland and Lecturer at the Center for the Law of Life Sciences, Faculty of Law, University of Bonn, Bonn, Germany

*E-mail: claudia.seitz@ugent.be, claudia.seitz@seitzriemer.com

Abstract. Advances in sciences—both in molecular biology and information technology—have enabled new understandings of genetic resources and biological processes which lead to an ongoing fundamental paradigm shift. The phenomenon of big data resulting from genomic sequencing has emerged as a consequence of next generation sequencing technologies that generate large data sets in biotechnological research. Technical progress in such big data generation and management has opened previously unprecedented possibilities for research, the development of new products and leads to a vast amount of data being generated at an unprecedented speed, ranging from the extraction of information to data analysis and interpretation. While there will always remain a reliance on genetic materials, one may observe a trend that research and development activities using genetic materials are increasingly supplemented or substituted by computerized research activities based on digital sequence information (DSI). These developments lead to a revolutionary transformation in the use of genetic resources (GR), which is currently undergoing radical changes and may be considered as a “fourth industrial revolution”. It results not only in exponential growth of generated genetic data, which may lead to innovation and new products and services, but also poses a range of new regulatory and legal questions due to its specific characteristics. The rise of big data raises specific legal challenges in terms of data ownership and intellectual property, data stewardship and governance, as well as technology transfer and licensing. One example is the question on how to regulate access and benefit-sharing for plant genetic resources and regulatory approvals for gene-edited plants as a result of new genomic technologies, which illustrates the use of DSI as an essential tool for research and development along with the requirements to exchange such information. Existing exchange mechanisms include, for instance, open-access databases and DSI platforms. Their data access and exchange policies would presumably be intended to maintain consistency with the objectives, policies and regulations of the Access and Benefit-Sharing system (ABS system). A large number of these questions are currently subject to discussions on an international, regional, as well as on a national level. This paper shall shed light on some of these questions to reflect the current and ongoing discussion on the issue and focus on the protection of DSI and address the question of the application of ABS systems to DSI.

Keywords: access and benefit-sharing, digital sequence information, genetic resources, genome sequencing, intellectual property.



1. Introduction

Innovation in plant breeding relies on specific knowledge, the development and application of new technologies, access to GR, and funding to exploit all those factors. Progress in molecular biology has led to the science of genomics and resulted in new techniques such as gene editing, synthetic biology, bio-nanotechnology and other evolving technologies, which can be considered as transformative tools for biological systems. The techniques may be used to enhance production in the pharmaceutical, agricultural, food and biotechnology sectors, to develop new diagnosis methods, personalized medicine, and new breeding techniques to improve plant varieties and animal breeds. In addition, in the last decades digitalization has become a major trend in the sector of GR. New characterization techniques for GR and new capabilities for collections of bioinformatics data, which resulted from innovation in bioinformatics, the use of artificial intelligence and the use of digital trust technologies, have become possible at reduced cost and have already created rapid scientific and technological progress.

The combination of both innovations in molecular biology and bioinformatics has led to new possibilities but also to new legal questions. The sequencing of genetic materials and the DSI of GR have raised a variety of questions [1]. To answer them, the terminology and different types of DSI on GR need to be defined first. Besides questions regarding potential implications of the use of DSI on GR for the conservation of biological diversity and their sustainable use and for the fair and equitable sharing of the benefits arising out of the utilization of GR, there are also several other practical legal questions. Especially the question of protecting and sharing of DSI and the resulting tension is not yet answered. The present paper shall therefore address and highlight a few remarks on the definition of DSI, its protection under the current IP system and the scope of application of the current ABS systems for sharing DSI.

2. Terminology and scope of digital sequence information

The transition from biological material to sequence data has led to intensive discussions on international level concerning the question how to address sequence data, and whether the regulations on GR concerning access, sustainable uses, as well as fair and equitable sharing of benefits arising from the use of GR shall also apply to sequence data [2]. These discussions are simultaneously taking place in multiple fora, such as the Convention on Biological Diversity [3] and its Nagoya Protocol of 2010 [4], in the FAO International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA) [5], in the World Health Organization (WHO) on the Pandemic Influenza Preparedness Framework [6], which not only uses sequence data but also other information connected with genes and related information, as well as in the UNESCO International Bioethics Committee [7].

So far, there is no precise terminology of DSI. The Conference of the Parties to the CBD in 2016 in Cancun adopted a decision that already incorporated the term DSI, but left the clarification of the terminology as a subject of further discussion in a study and an Ad Hoc Technical Expert Group (AHTEG) [2]. The term DSI is currently undefined, still not used in a uniform way, and interpreted differently by different stakeholders. In a narrow definition, it covers only sequence data. However, according to a broader definition, it seems to include other related information such as annotations and interpretation of data. In its broadest interpretation, DSI could comprise all immaterial, electronically saved data on GR, whereas GR is defined according to the Art. 2 of the CBD as genetic materials of actual or potential value, and genetic material is defined as any material of plant, animal, microbial or other origin containing functional units of heredity. The term genetic refers to the physical carrier of hereditary information, i.e. the deoxyribonucleic acid (DNA) and the ribonucleic acid (RNA).

It is not yet clear whether the term DSI should only apply to DNA or RNA sequences or whether it should be interpreted in a broader way to include protein sequences as well. The term sequencing could be defined as the process of determining and documenting the order of nucleotides or nucleobases on a given fragment of DNA or RNA, which are the building blocks of the chromosomes of organisms [8]. The lack of clarity in the definition of DSI causes problems in the current discussion since the subject matter is not clear, which causes a lack of common understanding and legal

uncertainties. Thus, the terminology and the scope of information which shall be covered by this definition is currently subject to a debate on an international level.

At the Fourteenth Conference of the Parties to the CBD, held in November 2018 in Egypt, an AHTEG was granted an extension to continue work on DSI of GR resources under the CBD and the Nagoya Protocol in the context of a post-2020 global biodiversity framework [9]. In a decision, the Conference of the Parties (COP) recognized that access to and use of DSI contributes to scientific research, and that many countries need further capacity to access, use, generate and analyze DSI. In addition, the COP noted that some parties have formed domestic measures that regulate the access to and the use of DSI as part of their ABS frameworks, but the parties have divergent views regarding the benefit-sharing from the use of DSI. Therefore, the COP had decided (i) to establish a science and policy-based process, by inviting submissions from governments and others regarding their views and information to clarify the concept (including relevant terminology and scope) and how their domestic ABS measures consider DSI, and the benefit-sharing arrangements from commercial and non-commercial use of DSI; (ii) to invite submissions from governments and others regarding information on capacity-building needs; (iii) to establish an extended AHTEG. Finally, the COP requested the Secretariat to synthesize the received submissions, and commission studies on the concept and scope of DSI, ongoing developments in the topic of traceability, existence of public and (if possible) private databases of DSI, and how existing domestic ABS measures address benefit-sharing arising from commercial and non-commercial use of DSI. The AHTEG was given the task to deliberate the synthesis of views, develop options for operational terms and their implications to formulate conceptual clarity, identify important areas for capacity building and submit all the outcomes to be considered by the working group on the post-2020 framework.

3. Digital sequence information and intellectual property

In the context of intellectual property (IP) protection of DSI, patents, copyright and trade secret protection raise the most interesting IP protection issues as the full eligibility of DSI for coverage under either regime is currently neither clear nor free from doubt. Genetic sequencing and DSI as such are generally not an invention. Thus, under most patent laws and as a general principle, mere products of nature and discoveries, where no human innovation is involved, are not patentable. Since the decision of the US Supreme Court in the *Myriad* case in 2013 [10], it is now widely acknowledged that the mere identification of a genetic sequence as such is not patentable since it is a discovery and not an invention. In addition, such scientific discovery is, by definition, universal and must remain accessible to everyone.

In the European Union (EU), however, the EU Biotechnology Directive [11] of 1998, does not generally exclude patents on gene sequences isolated from nature. According to Art. 3.2 of the Directive, the definition of biological material includes any material containing genetic information and capable of reproducing itself or being reproduced in a biological system. Art. 5.2 constitutes that elements isolated from the human body, including sequences and partial sequences of a gene, may constitute a patentable invention, even if the structure of that element is identical to that of a natural element. Art. 1.1. and 3.1. of the EU Directive constitutes, however, that a patent is granted for an invention which is new, involve an inventive step and which is susceptible of industrial application. Thus, a patent under the EU Directive still requires an invention as well.

Moreover, copyright protection of DNA sequence information seems to be a weak protection. Although this information may be copyrightable as a “work”, if it is original, in principle there is no specific protection in place. An analogy to the copyright of the source code of computer software programs can be drawn, since the algorithms which constitute the source code can be protected under copyright laws. There remains a question, however, of what extent a DNA sequence is comparable to an algorithm of a software program. The mere detection of sequence information may not be comparable, whereas the encoding functions of DNA sequences and their recombination within the scope of synthetic biology could be considered as similar to algorithms of software programs since this constitutes not a mere detection but a new and human-designed combination of sequences to

create something new and innovative. If the focus is on the codes of DNA sequences, it could be argued that sequence information could fall under the category of literary works since it could be compared to words, characters, numbers, symbols or signs that express something specific like a specific kind of language. However, if the DNA sequence is just the expression of a natural functionality, it seems to be difficult to apply copyright laws to this information, which is similar to literary works. In addition, there is neither case law on copyright protection of DNA sequences so far nor an existing copyright law which explicitly mentions genomic data, DSI or any other form of data of genetic resources.

According to the “WIPO Guide in IP Issues in ABS-agreements”, copyright protection may arise when advanced characterization data about GR are created, such as DSI. However, whether copyright applies is subject to applicable law [12]. In addition, the guide explains that existing fact-finding and scoping studies have identified the ways in which copyright and other IP are asserted for sequence information and its applications, in various sectors and under different scenarios, together with the implications for ABS, including monitoring, as one important area among several which needed further and deeper investigation.

Since patent and copyright laws seem not to cover DSI at all or at least sufficiently, one could think about know-how protection of sequence data as a kind of a trade secret. In this case, however, special precautionary measures need to be in place in order to protect the secrecy of the information. If the data, however, is made publicly available in a public database or platform, then there is no precautionary measure in place and thus no secret information which could be kept protected.

As a result, it is clear that the current IP system does not cover DSI explicitly. On the other hand, the importance of DSI in the innovation process and in the management and production of GR-based products is evident and increasingly growing, and the results of the innovation process could lead to claims of ownership of relevant processes and products. The background of this legal situation is the different role of IP on the one side and GR and DSI on the other: The objective of IP is to protect innovation and creativity, whereas the objective of GR policy and DSI is access and benefit-sharing (according to the Nagoya Protocol). DSI from GR and research may lead to IP rights and there are multiple interfaces between to two sides, but the underlying objectives for IP is protection and thus their regulations are different.

4. Access and benefit-sharing

Recently, a debate has emerged whether to apply existing regulations on access to GR resulting from the Nagoya Protocol and its implementation on regional and national level to DSI and, if so, to what extent [2,7,13,14]. Besides the unclear terminology and scope of DSI, as already explained, the questions that are currently discussed in this context are whether the existing ABS system applies to DSI: if so, to what extent the system can be applied; and if not, whether a completely different ABS system modelled on Art. 10 Nagoya Protocol is needed to sufficiently and appropriately address the transformation from biological material to sequence data in order to meet the underlying public policy requirements of the CBD and the Nagoya Protocol. This discussion on whether and how DSI is covered by the Nagoya Protocol is of fundamental importance for the ABS system, since DSI enables the access to genetic data independent from the access to the biological material for GR research. This leads to the consequence that there is no need for travelling to a country or entity that provide a particular GR, to start negotiations with this country under the ABS system, and to sign appropriate agreements if the coded characteristics of the GR can be downloaded from an electronic database or portal of DSI.

Currently, in most jurisdictions, sequence information downloaded from public databases is not covered by the Nagoya Protocol. However, some countries such as Brazil have their own national legislation for access to GR, which also covers rights over DSI. While sequence information can be downloaded from public databases without any obligations by the ABS system, some argue that DSI is already covered by the Nagoya Protocol or that its provisions should be adapted to cover sequence information. This current debate is leading to an unclear legal position and causes legal uncertainty.

The negotiation history of the Nagoya Protocol and its objectives make it seem clear without a doubt that DSI is the result of sequencing activities on GR, and thus should be considered as the results of research and development activities or as a form of utilization of GR according to Art. 2 of the Nagoya Protocol. Yet, it is still unclear to what extent the ABS system applies to DSI [15].

5. Concluding remarks

The fundamental and radical transition from material to data is unique in history. It brings along a lot of changes and possibilities to the global research and use of genetic materials. However, the technological use of DSI in a rapidly changing scientific environment requires solutions and answers to legal, policy and scientific implications of regulating DSI, which is becoming an increasingly complex challenge. Whereas one could argue that genetic sequences do not fall into the scope of the Nagoya Protocol, since free sharing of DSI is a common scientific practice and that any other solution would be counterproductive for global research or even dangerous for public health if applied to the pharmaceutical sector, the answer is not that easy. There is a need for a common terminology for DSI on an international, regional and national level, and what should be covered by this term. Also, there should be a discussion on the question whether a specific (IP) protection *sui generis* should be established to protect DSI. New innovations may create new needs for protection through new uses of existing IP law, IP-related technological protection measures for GR and DSI, or through similar *sui generis* protection rights. Finally, there should be an open discussion on the question whether provisions of the Nagoya Protocol should be amended to address the regulation of DSI and to clarify the questions regarding the applicability of the current ABS system on DSI. The possibility of the production of functional biologics which are produced automatically from digitally transmitted DNA sequences, has the potential to replace biological material and will increasingly enhance research and development processes in a wide range of sectors. In an age of synthetic biology, when an organism can theoretically be made artificially using genetic sequencing or downloading sequence information from a public database, questions arise about the importance of biological material and the functionality of the current ABS system.

Although it is currently not yet possible to synthesize more complex organisms such as crops from scratch, the increasing importance of the replacement of biological material by data is evident. By combining sequence data with new and innovative gene editing technologies, such as editing gene using CRISPR/Cas9, the importance of sequence data is obvious since a specific DNA sequence can be used and introduced into another organism without access to the biological material, where the specific DNA information came from. Presently, this can be done without negotiating any material transfer agreement under the current ABS system.

6. References

- [1] Reichman J, Uhler P and Dedeurwaerdere T 2016 *Governing Digitally Integrated Genetic Resources, Data, and Literature: Global Intellectual Property Strategies for a Redesigned Microbial Research Commons* (New York: Cambridge University Press)
- [2] CBD 2016 *Decision Adopted by the Conference of the Parties to the Convention on Biological Diversity* (Cancun)
- [3] CBD 1992 *Convention on Biological Diversity* (Rio de Janeiro)
- [4] Secretariat of the Convention on Biological Diversity 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* (Canada: United Nations Environmental Programme)
- [5] FAO 2009 *International Treaty on Plant Genetic Resources for Food and Agriculture* (Rome: Food and Agriculture Organization)
- [6] World Health Organization 2016 *Pandemic Influenza Preparedness Framework* (Switzerland)
- [7] International Bioethics Committee 2016 *Preliminary Report of the IBC on Big Data and Health* (Paris)

- [8] Sollberger K 2018 *Digital Sequence Information and the Nagoya Protocol* (Australia)
- [9] Tsioumani E, Appleton A, Finnegan L, Rude A, Schabus N and Tsioumanis A 2018 Summary of the UN biodiversity conference *ENB* **9** 1–27
- [10] United States 2013 *Supreme Court of the United States: Association for Molecular Pathology v. Myriad Genetics, Inc.* **569**
- [11] European Union 1998 Directive 98/44/EC of the European Parliament and of the Council of 6 July 1998 on the legal protection of biotechnological inventions *OJ L* **213** 13–21
- [12] ABS Capacity Development Initiative 2018 *A Guide to Intellectual Property Issues in Access and Benefit-Sharing Agreements* (Switzerland: World Intellectual Property Organization)
- [13] Karger E 2018 *Study on the Use of Digital Sequence Information on Genetic Resources in Germany*
- [14] TB Spranger 2017 *Expert Opinion on the Applicability of the Convention on Biological Diversity and the Nagoya Protocol to Digital Sequence Information* (Bonn: Federal Ministry of Education and Research)
- [15] Secretariat of the Convention on and Biological Diversity 2010 Uses of genetic resources *Convention on Biological Diversity-ABS* 1–8