UK BioIndustry Association

The UK BioIndustry Association (BIA)'s response to the CBD Secretariat's consultation on the impact of Digital Sequence Information (DSI) regulation in the Nagoya Protocol

About the BIA

- Established in 1989, the BioIndustry Association (BIA) is the United Kingdom (UK) trade association for innovative bioscience enterprises. The BIA represents over 300 member companies, including innovative start-ups and small and medium-sized enterprise (SME)[1] bioscience companies, academic research and philanthropic organisations, and service providers to the UK bioscience sector;
- Our members are responsible for over 90% of biotechnology-derived medicines currently in clinical development in the UK and are at the forefront of innovative scientific developments targeting areas of unmet medical need;
- Many of our members are pre-revenue SMEs operating at the translation interface between academia and commercialisation;
- Our goal is to promote innovative bioscience in the UK for the benefit of global public health by enabling our world-leading research base to deliver healthcare solutions that can truly make a difference to people's lives.

Executive summary

- The BIA and its members support the objectives of the Convention of Biological Diversity (CBD) and the Nagoya Protocol (NP), however, we have concerns about the inclusion of Digital Sequence Information (DSI);
- There are vast amounts of DSI in publicly available databases. According to one estimation, there will soon be quintillions (>10^{18}) of nucleotides of DNA sequence data publicly available;
- Almost all bioscience companies and researchers rely heavily on DSI to conduct research. There are many different types of DSI used in R&D for various reasons. These include, for example, the sequencing of biomedical samples from patients to identify a harmful pathogen or environmental samples to assess what species are present;
- There have been concerns voiced over the access and benefit sharing (ABS) objective of the NP and suggestions that the current ABS system is not working. However, there is no evidence to suggest that the incorporation of DSI will help achieve the ABS objective nor remedy the ABS system. If there are shortcomings with the ABS system, it is largely due to the lack of provider country laws which facilitate access and thus generate benefits;

[1] Small and medium-sized enterprises (SMEs) are non-subsidiary, independent firms which employ fewer than a given number of employees. This number varies across countries. The most frequent upper limit designating an SME is 250 employees, as in the European Union. However, some countries set the limit at 200 employees, while the United States considers SMEs to include firms with fewer than 500 employees. See https://stats.oecd.org/glossary/detail.asp?ID=3123

- The BIA has consulted its members to gather their views and examples of how they might be impacted by the proposed incorporation of DSI into the NP. **All responding members strongly disagree with the proposed incorporation of DSI into the NP.** There are three main reasons for this disagreement:
    1. **The incorporation of DSI into the NP will lead to further legal uncertainty and compliance difficulties for SMEs**
    2. **DSI regulation in the Protocol will hinder SMEs' R&D in all biological fields, including medical, agricultural, and industrial biotechnologies**
    3. **The incorporation of DSI into the NP poses serious public health concerns**

A. Use of DSI in the modern bioscience sector

i. Importance and ubiquity of DSI

DSI has been generated, stored and used for several decades in vast and increasing quantities. It is estimated that "publicly available databanks contain now contain quadrillions (>10^{15}) of nucleotides of DNA sequence data, soon to be quintillions (>10^{18} bases). These have been collected from over 300,000 different species of organisms."[2] Researchers often analyse datasets that are many terabytes in size.[3] This has created new challenges in acquiring, analysing, storing, and distributing such data. Any DSI regulation would have to carefully consider these challenges, the vast amounts of data publicly available, and the benefits this extensive information can bring to all of humanity.

ii. Different types of DSI

There are many biological sequence databases. Taking one as an example, the National Center for Biotechnology Information (NCBI) sequence database is designed around a model of biological sequence data that classes sequences according to types (nucleic acid, which may be further sub-categorised as DNA or RNA, and protein) and allows many other attributes to be recorded.[4]

Researchers use DSI for many different reasons and to generate various outputs. As a result, researchers use DNA sequence data with different qualities, such as:

- DNA 'barcodes' - short stretches of DNA with sequences that are conserved enough to find and yet diverse enough to allow researchers to identify what organism they are from
- Gene sequences - sequences that include the start and stop instructions and all the necessary DNA codons to create a gene product e.g. a protein
- Regulatory DNA - stretches of DNA that do not code for proteins but instead have effects on e.g. the processing of genes

[2] Pevsner, J. (2015). *Bioinformatics and functional genomics*. Chichester, West Sussex: Wiley Blackwell.
[3] Ibid.
[4] Full details available at https://www.ncbi.nlm.nih.gov/IEB/ToolBox/SDKDOCS/DATAMODL.HTML

- Whole genomes - the complete genome sequence of an organism. As each organism is unique, there may be standard or consensus (compiled from multiple, even thousands) genome sequences.

The above list demonstrates that the term “digital sequence information” is broad in scope and does not encompass a single type of data. In considering whether DSI should fall under the scope of the NP, it will be vital to carefully consider and consult with researchers about which types of sequence data (or data used in which ways) would be considered within scope.

iii. Uses of DSI

There are many reasons why researchers might sequence genetic data. The list below is non-exhaustive but illustrates some reasons for determining and using DSI:

- Biomedical samples from patients to identify the pathogen
- Pathogens over time and space to investigate epidemiological spread
- Environmental samples to assess what species are present
- Organisms to infer their evolutionary relationships
- Pathogens to identify potential targets to develop therapeutics or vaccines against
- Molecular biology lab work to check a result is as expected. (In molecular biology, researchers routinely attempt to create genetic tools to help with either basic or applied research. These tools very often for example include cloning some sequences (e.g. a gene(s) and/or a promoter and/or other elements) into other sequences (e.g. a DNA vector 'backbone'). DSI would be used first in designing this construct, and later generated in sequencing to verify its correct construction).

B. Negative consequences of incorporating DSI into the Nagoya Protocol

The BIA has consulted its members to gather their views and examples of how they might be impacted by the proposed incorporation of DSI into the NP. All responding members strongly disagree with the proposed incorporation of DSI into the Protocol. This section explains the three main reasons for this disagreement.

First, however, it must be noted that any decision to amend the scope of the NP must be based on clear evidence that the ABS objective is not working, the current ABS system is failing, and that the incorporation of DSI would help achieve the ABS objective and remedy the ABS system. Currently there is no such evidence.

If there are shortcomings with the ABS system, it is largely due to the lack of provider country laws which facilitate access and thus generate benefits. Once addressed, and comprehensive legal frameworks of national ABS laws are put in place, concerns about the lack of benefit sharing related to genetic resources access and use should disappear.

1. The incorporation of DSI into the Protocol will lead to further legal uncertainty and compliance difficulties for SMEs

Many BIA members are already experiencing numerous legal and practical uncertainties with the NP and the EU Regulation 511/2014. The possible incorporation of DSI into the NP would exacerbate these uncertainties. The uncertainties relate to the generation and utilisation of DSI and would particularly affect SMEs, which lack the resources to ensure due diligence requirements are met.

Examples of these uncertainties include:

i. Determining the country of origin of a sequence

Digital sequences are frequently available in public databases. The countries of origin of these sequences are often unknown. This was acknowledged by the US Council for Trade-Related Aspects of Intellectual Property Rights: "Even within publicly accessible international and national gene banks, there are many resources where the country of origin is unknown."[5]

Research and other uses of DSI often involves the combination, editing, and refinement of potentially large amounts of information originating from many sources.

Furthermore, a researcher who wanted to access a particular digital sequence could search for equivalent files from different countries of origin, and access the sequence from the country with the least burdensome (or no) ABS regulation.

This raises several questions:

- How will the countries of origin of sequences be determined when these are not known?
- How will the countries of origin of combined, edited, and/or refined sequences be determined?
- If an organism exists in multiple countries and a sequence based on the organism is utilised, how will it be decided what country should benefit?

ii. Downstream use of digital sequences

With digital sequences readily available in public databases, several questions are raised regarding the downstream use of these sequences.

- How would the downstream use of digital sequences be regulated without placing undue burdensome due diligence requirements on researchers, in both commercial and academic institutions?

[5] Council for Trade-Related Aspects of Intellectual Property Rights, Communication by the United States, Article 27.3(b), Relationship Between the TRIPS Agreement and the CBD, and the Protection of Traditional Knowledge and Folklore, IP/C/W/469, para. 37 (Mar. 13, 2006), cited in "The Nagoya Protocol and Synthetic Biology Research: A Look at the Potential Impacts', The Wilson Center, 2013, p.14: http://www.synbioproject.org/site/assets/files/1291/nagoya_final-1.pdf.

- If a digital sequence is uploaded to a database in accordance with the local ABS regulations, will subsequent users of that sequence also be required to comply with the same ABS regulations?
- If a digital sequence is uploaded without the proper permission from the country of origin, will subsequent users be held legally responsible if they utilise the sequence?

iii. Accidental creation of sequence similarity

The design of novel sequences on computers is now an everyday practice in laboratories around the world. It is entirely plausible that a researcher could accidentally create a digital sequence identical to one found in nature.

- How could the researcher prove that the sequence was created by accident?
- What due diligence obligations would apply and how would these be policed?

iv. Synthetic sequences

Some sequences, such as Venter's *Mycoplasma laboratorium* genome, are entirely synthetic. Any DSI regulation would have to consider what due diligence requirements would apply to researchers working with entirely synthetic sequences.

The regulatory framework answering each of these questions would be complicated to develop, manage, and police. The regulatory framework would also place burdensome due diligence requirement on bioscience companies. In particular, the many legal and practical uncertainties posed by the possible regulation of DSI would have a hugely negative impact on scarcely-resourced SMEs involved in R&D and public health activities.

Responding to the BIA's consultation, our members wrote:

> "The only likely outcomes of the incorporation of DSI in the Nagoya Protocol would be the occasional criminalisation of companies and individuals who are not operating any differently to the rest of the biotechnology industry but are simply unlucky. It has no possibility of benefitting developing countries."
>
> "We use sequences from genetic databases where the source of the original material is often not defined. For example, a gene from a bacterium that causes bacterial diarrhoea may be from a developed or a developing country source. If this is incorporated into a vaccine, there is no way of knowing the country of origin. If the origin of DSI was not originally available but is later proven, we risk conviction even if we originally tried diligently to identify the source."
>
> "We occasionally assemble new genes by producing alignments of original genetic sequences from sources obtained worldwide, to generate consensus sequences that compensate for natural variation – this is particularly useful for vaccine applications. A novel sequence that is an amalgam of other sequences cannot be treated as the source of a single country, and identifying the origins of the component sequences would be very difficult or impossible, as many sequences can be used to compile such a consensus."

> "We frequently use codon optimisation to change the original sequence to one that expresses better in particular bacteria. If the final sequence differs from the original sequence, it cannot be regarded solely as the genetic property of an originator country."

2. DSI regulation in the NP will hinder SMEs' R&D

The incorporation of DSI in the Protocol will restrict both access to and the use of digital genetic information which is in the public domain. As demonstrated in Section A above, this information is a crucial tool in biomedical research. Its restriction by way of complicated and burdensome due diligence requirements to ensure compliance will adversely affect R&D in non-commercial and commercial sectors and may delay, or even prevent, new medicines vital to public health from reaching patients.

Responding to the BIA's consultation, our members wrote:

> "The additional level of delays and bureaucracy that will result will make it less likely that medicines that would specifically benefit developing countries would be developed in the first place, with vaccines being particularly vulnerable in this respect."

> "We see availability of these publicly available sequences to research and commercial scientists as being crucial for supporting scientific research. For example, these databases and publications are widely used in the development of new diagnostic tests against pests and diseases or, in the development of novel reagents used in molecular biology."

3. The incorporation of DSI into the NP poses serious public health concerns

The increased legal uncertainty for life science companies and the impact of DSI regulation on R&D would inevitably affect public health. While DSI regulation would impact all medicines development, its implications would be most evident in rapid response R&D, vaccine production, and in the global effort against anti-microbial resistance (AMR).

Virus epidemics, such as Ebola and Zika, need to be sequenced and shared quickly due to the viruses' rapid evolutionary rates. The below quote from a peer-reviewed scientific journal demonstrates the importance of the fast sequencing and sharing of Ebola virus genome:

> Genome sequencing provides a high-resolution view of pathogen evolution and is increasingly sought after for outbreak surveillance. Sequence data may be used to guide control measures, but only if the results are generated quickly enough to inform interventions.[6]

[6] Quick, J. et al. "Real-time, portable genome sequencing for Ebola surveillance", Nature 530, Feb. 2016, pp. 228-232. http://www.nature.com/nature/journal/v530/n7589/full/nature16996.html?foxtrotcallback=true

The incorporation of DSI into the NP risks preventing researchers from the quick sharing of sequences necessary to develop vaccines to combat viruses such as Ebola.

DSI regulation would also undermine initiatives such as GISAID, which promotes the international sharing of all influenza virus sequences, related clinical and epidemiological data associated with human viruses, and geographical as well as species-specific data associated with avian and other animal viruses, to help researchers understand how the viruses evolve, spread and potentially become pandemics.[7] Unless GISAID were explicitly given the status as a recognised international instrument, DSI regulation would hinder the sharing of sequences and the public health benefits arising therefrom.

Responding to the BIA's consultation, our members wrote:

> "The rapid sharing of sequence data is essential in the fight against infectious disease outbreaks, the monitoring of drug resistance and other matters of international public health importance. There may, justifiably, be a concern that the Protocol might introduce regulatory hurdles that could inhibit an ability to take part in national or international collaborative research efforts to address these issues."
>
> "The additional level of delays and bureaucracy that will result will make it less likely that medicines that would specifically benefit developing countries would be developed in the first place, with vaccines being particularly vulnerable in this respect."

[7] https://www.gisaid.org/