International Chamber of Commerce
Policy and Business Practices

DIGITAL SEQUENCE INFORMATION AND THE NAGOYA PROTOCOL

SUBMISSION TO THE CBD

Prepared by the ICC Task Force on Access and Benefit Sharing

Summary and highlights

INTERNATIONAL CHAMBER OF COMMERCE
The world business organization

450/1111
DYE/abs 14.09.2017

ICC and its members support the objectives of the Convention on Biological Diversity (CBD) and the Nagoya Protocol (Protocol).

ICC is very concerned about the international discussions proposing to expand their scope to include digital sequence information (DSI) or its use. DSI is not within the scope of the CBD or the Protocol, because the definition of genetic resources relates to genetic material and not abstract information. Expansion of the scope to include DSI would challenge the agreement embodied by the Protocol and amount to its total re-negotiation. Inclusion of DSI into the scope of the CBD and the Nagoya Protocol would also hinder rather than help achieve the objectives of the CBD and the Protocol.

Therefore, ICC is strongly opposed to the expansion of the scope of the CBD and the Protocol and recommends giving priority to their effective implementation in order to achieve their objectives.

Introduction

ICC supports the objectives of the Convention on Biological Diversity (CBD) and the Nagoya Protocol (Protocol), and has actively contributed to discussions on the development of the Protocol and its translation into national and international legislation.

ICC is very concerned about the current international discussions proposing to expand the scope of the CBD and the Protocol to include digital sequence information (DSI) as such or its use. DSI and its use are not within the scope of the CBD or the Protocol. The increasing use of DSI to advance the conservation and sustainable use of biodiversity does however make a significant contribution to achieving the objectives of both the CBD and the Protocol. ICC is therefore of the opinion that creating new rules related to access to and use of DSI will seriously impede research and development, and as a consequence will hinder rather than help achieve the objectives of the CBD and the Protocol. As a way forward, priority should be given to the effective implementation of the Protocol based on its agreed scope and to putting in place national rules and infrastructure to safeguard the sustainable use of genetic resources and provide legal certainty for their access and use.

The scope of the Convention on Biological Diversity and the Nagoya Protocol

An expansion of the agreed scope to include DSI or its use would change a fundamental aspect of both the CBD and the Protocol, and would call into question the hard-won consensus which they represent.

The definition of a “genetic resource” - as provided by Article 2 of the CBD and referred to in Article 2 of the Protocol - is “*genetic material of actual or potential value*”. “Genetic material” is defined as ”*material of biological origin containing functional units of heredity*,” with genes recognised as the basic units of heredity. It follows from this definition that, in the absence of material, the resource in question does not qualify as a genetic resource under the CBD or the Protocol. Genetic resources are therefore understood to cover *materials* such as organisms, or parts thereof, in which genetic

material is present (with the exception of material of human origin which has been explicitly excluded). In other words, the term refers to tangible genetic material which must physically contain genes. It therefore follows that intangible DSI as such cannot constitute a genetic resource as defined by the CBD.

Further expanding the definitions of “genetic resources” in the CBD and/or the “utilisation of genetic resources” in the Protocol to include DSI or its use would create legal uncertainty around the use of such information and as to how access and benefit obligations would apply. This would be contrary to the basic understanding of the aims of the Protocol negotiations, namely on the one hand to create legal certainty for access to genetic resources and on the other hand to provide for compliance mechanisms based on such legal certainty. This issue will be further exacerbated by the technical difficulty or material impossibility in many cases of establishing which species a genetic sequence comes from since genetic sequences can be common to several species.

Negative implications on the objectives of the CBD and the Protocol

The three objectives of the CBD are as follows: 1) conservation of biological diversity, 2) sustainable use of its components, and 3) appropriate access to genetic resources and the fair and equitable sharing of benefits arising out of their utilisation. The Protocol’s key objective is to provide a transparent legal framework for access to genetic resources and the fair and equitable sharing of benefits arising out of the utilisation of genetic resources in order to encourage the sustainable use and conservation of biodiversity.

It needs to be stressed that it is of paramount importance that DSI now in the public domain continues to remain freely accessible to achieve the broader aims of the CBD. This is consistent with the requirements of the CBD that include: promoting and encouraging research that contributes to the conservation and sustainable use of biological diversity (Article 12); public awareness and education (Article 13); facilitating access to genetic resources (Article 15); facilitating access to or transfer of technology (Article 16); exchanging information, including scientific research (Article 17); technical and scientific cooperation (Article 18); and promoting and advancing the benefits of biotechnology (Article 19). Open exchange of scientific information, including DSI, contributes to these activities that support the objectives of the CBD and the Protocol, and should be explicitly qualified as benefit sharing in itself. Imposing further obligations on this type of data would go against the objectives of the CBD and Protocol as set out above. Cooperation and the open exchange of information are also consistent with established principles of ethical and responsible scientific research.

Negative implications on research and development

Effective and efficient research and development is critical to achieving the objectives of the CBD and the Protocol. The current discussions in the context of the Protocol are centered around cases with a one-on-one relationship between the accessed genetic resource and the created product or benefit - for example, the isolation of a new chemical entity from an exotic plant, which is further developed into a commercial product. In these cases, identifying and applying ABS obligations is already complex. This complexity is markedly higher in state-of-the-art research and development, e.g. in the biotech sector, where in principle multiple genetic resources are accessed from different sources and

used for the development of one commercial product. This complexity is even more present when DSI is involved as thousands of sequences might be used in the development of one product.

In order to provide some necessary insights into the complexity of research and development, especially in relation to DSI, below is a case study from the biotech sector describing the development of a protein-engineered enzyme aimed at improving the nutritional value of animal feed. This case is representative of today's application of biotechnology in research and development in various sectors. The detailed case study is attached as Annex I. The case clearly demonstrates that in state-of-the-art bioinformatics projects, hundreds to thousands of (amino acid or nucleic acid) sequences may be used to develop a particular commercial product. The final product has a sequence that represents an "average" of all input sequences; as a consequence, it is virtually impossible to determine the relative value of each individual input sequence. In addition, should DSI be included in the scope of the Protocol, the administrative burden of negotiating a myriad of ABS agreements for sequences with debatable input value will be significant. The case also shows that while natural biodiversity may provide inspiration, value is actually created by computational and experimental means.

The case study demonstrates the practical difficulties and legal uncertainty that would arise with the expansion of the scope of the CBD and the Protocol by including DSI or its use. Furthermore, since most of the DSI contained in gene banks or databases are public and freely available for any potential research and development activity, imposing new (ABS) requirements on their use would also amount to a retroactive application of ABS obligations and add further uncertainty, as well as create an additional barrier for the use of such DSI. This legal uncertainty would negatively affect research, education, and scientific innovation related to the use - and a fortiori the conservation - of genetic resources, with social costs that are far reaching and underestimated in current discussions under the CBD and the Protocol.

Sharing of information in the public domain is thus foundational for the advancement of science and innovation around the world, and for the achievement of the key objectives of the CBD and the Protocol, especially the sustainable use of genetic resources and the conservation of biological diversity. A recent article in Science also points to the fact that "*benefit sharing rules for digital DNA would be hard to enforce, and could clash with the open-access culture of research*". It has furthermore been stated that "*there is also a concern that if DSI was to be added to the remit of the Nagoya Protocol, the official process to request access to the digital sequence data (which could no longer be open access) and negotiations focused in adequate benefit sharing could discourage or unnecessarily prolong the research process*"[1]. Indeed, a system which imposes ABS requirements on DSI would be unworkable for both users and government authorities because compliance, monitoring and checking would be extremely burdensome or even impossible to achieve.

The entities active in research with genetic resources that would be most negatively affected would be academic and other public institutions which are often involved in the first steps of the R&D process. This system of open innovation works with subsequent users adding incremental value (through data and knowledge) and relies on legal certainty and clarity on rights and obligations for all players in the R&D chain. The R&D process would be substantially undermined and become (materially and economically) impossible. This would have a chilling effect on research with genetic resources, would have a very negative effect on the likelihood of innovative products and would result in a very substantial loss of socio-economic benefits for society as a whole. It should be stressed that much is at stake for major international development collaborations (i.a. those targeting food security)

[1] Servik, K (2016) Rise of digital DNA raises fears of biopiracy, Science. DOI: 10.1126/science.aal0395.

given the significant investment in modern sequence-based technologies. The significant "non-monetary" benefits delivered to developing countries could be diminished in the process of trying to extract monetary benefits from DSI.

The way forward

Mutually agreed terms (MAT) between the provider and the user of genetic resources are of great importance as they are a key component of the recognition of the sovereign rights over genetic resources and the essentially bilateral nature of the Protocol.

Under the Protocol, establishing fair and equitable benefit sharing is based on MAT. MAT are used to bilaterally negotiate fair and equitable benefit sharing between a provider and user for the utilisation of the genetic resource and the use of the information obtained therefrom.

The CBD established an approach to sustainable use (Article 10), and minimum standards for accessing genetic resources (Article 15). Individual national governments were given the right to define requirements for MAT, and this principle was reaffirmed in the Protocol[2].

Instead of creating legal uncertainty by expanding the scope of the Protocol, priority should be given to the implementation of the Protocol to put into place the frameworks, infrastructure and practices that will allow its goals to be realised. Many national governments have invested considerable time and resources to implement the Protocol at the national level. Businesses around the world are also making significant efforts to understand national implementing measures and integrate them into their operations. Such efforts are premised on the principles and standards agreed in the Protocol, and shifting the goal posts at this stage will cause significant problems for both governments and users. It is in the best interest of parties developing new technologies to focus on implementing the CBD and the Protocol as written rather than attempting to renegotiate terms and concepts that have been discussed for over ten years and agreed when the Protocol was approved by the Parties in 2010.

ICC would also like to propose that an evaluation should be made of the impact of the Protocol to determine its effect on access, benefit creation and benefit sharing, as well as on conservation and sustainable use of biological diversity. Preliminary evidence suggests that the uncertainty created by the lack of clarity in the Protocol and its implementation is having a chilling effect on the sourcing and use of genetic resources for innovative activities. It is therefore essential to take stock of the effect of the implementation of the Protocol to date before considering whether to change its scope.

ICC and its members remain committed to the objectives of the CBD and the Protocol, which are also strongly embedded in the UN Sustainable Development Goals. The members of ICC are active players in the value chain involving genetic resources, and invest substantial resources in the R&D process to unlock the potential of genetic resources. We will continue to cooperate with all relevant stakeholders to ensure effective, workable and science-based solutions which create a proportionate regulatory environment to enable continued characterisation, conservation and sustainable use of genetic resources.

[2] Article 6.1, which established a requirement for PIC, may be in contradiction with Article 15.5, which allows governments to decide PIC is unnecessary. Article 6.3 of the Protocol may be interpreted as being consistent with the CBD.

ANNEX I: Case study

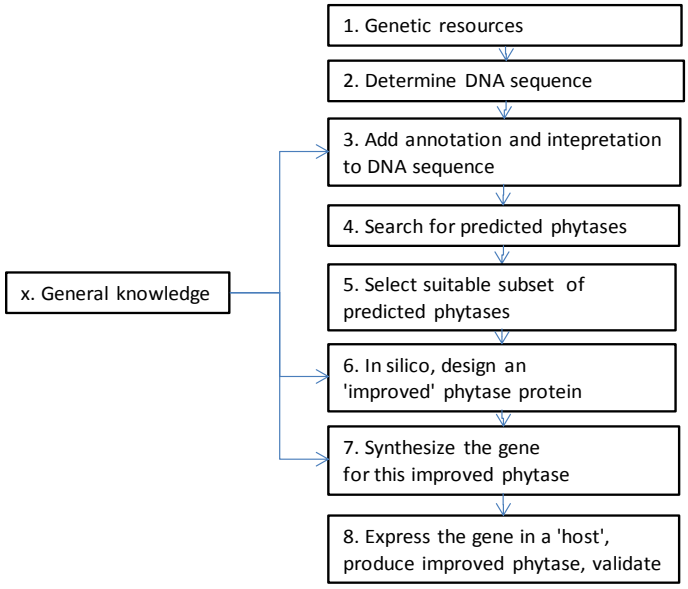
Development of a new consensus phytase to improve the nutritional value of animal feed

It has been demonstrated that a “consensus protein” has the potential to be more thermostable than any of the individual protein sequences used in its design. For the design of a consensus protein (e.g. a consensus phytase), a (large) set of homologous sequences for the protein of interest (e.g., all known microbial phytase amino acid sequences) are aligned, and for each amino acid position, the most frequent amino acid is calculated and selected. Next, this consensus amino acid sequence is converted into a corresponding DNA sequence, and this DNA sequence is then synthesized chemically.

The consensus protein is then expressed in a suitable host, and tested for its thermostability. If it is found to have the desired thermostability, the product will be developed for commercialisation.

This consensus approach has been applied to the design of many proteins and is nowadays an established protein design tool.

Scheme showing the steps in the case study:



Legend:

- ✓ Phytase is an enzyme used as a feed additive that ensures better nutrient uptake, for example in pigs.
- ✓ Step 1-3 are not targeted to any specific application, nor to a specific genetic resource, or species.
- ✓ General knowledge used for step 3 and 6 results from all genetic and biochemical analysis and protein structure on huge amounts of biological material, over the last decennia.
- ✓ General knowledge used for step 7 is based on genetic analysis on a huge amount of material of the future expression host.
- ✓ The product in step 8 requires ‘using’ the general knowledge (box x) at least at 3 steps, the subset of some 153 phytase sequences in step 4, (or the totality of 2494 or 1473 phytases in step 3).
- ✓ Similar processes are followed for laundry detergents, food enzymes, etc.

Detailed description of the case (for illustration purposes)

(1) Retrieve and select phytase amino acid sequences from public databases by applying a sequence similarity search

The initial commercial benchmark is taken as a starting point to find – through a homology search – related phytases in sequence databases (often public databases). Related phytases mean those with a similar (highly homologous) amino acid sequence. The higher the similarity (usually called homology and quantified as a percentage of identity), the more the phytases are related (homologous). Almost all protein sequences in databases (often public) are derived from nucleic acids sequencing. Those parts of the nucleic acid sequences that code for proteins (called open reading frames, or ORFs) are stored in protein sequence databases as amino acid sequences. As a protein's structural and functional properties are determined by its amino acid sequence, proteins are classified according to their amino acid sequence.

In this demonstration case, the amino acid sequence of the commercial benchmark was used as a bait, to find homologous phytase sequences in a protein sequence database at NCBI. Using a set threshold, 2494 protein sequences with a significant sequence similarity to the commercial benchmark were retrieved. To prevent a bias due to over-representation of almost identical sequences, the sequences were subjected to a clustering approach. This left 1473 sequences with a mutual sequence difference of at least 10%.

Next, the number of sequences was further reduced arbitrarily, by using 3-dimensional information from the protein database as a filter. Twelve amino acids considered to be important for catalysis were defined and only homologous amino acid sequences having all these 12 specific amino acids were retained, which reduced the number of homologous sequences from 1473 to 153.

(2) Determine the consensus sequence based on a sequence alignment of the pre-selected 153 homologous phytase amino acid sequences

To determine the consensus sequence at each position of the amino acid sequence, the most frequently occurring amino acid is calculated at that position. If, at selected amino acid sequence positions, two or more amino acids share the same highest frequency, one of these highest frequency amino acids is selected arbitrarily.

It must be emphasised that the consensus sequence is rather disconnected from the donor sequences. Since at every amino acid position, merely the most frequently occurring amino acid has been selected, it is impossible to trace back the origin of a particular amino acid to a specific 'parent' phytase.

Calculating the homology between the consensus phytase and each of the input phytases by performing an all-to-all pairwise alignment shows that, in general, they differ by more than 30%. This is very similar to the mutual difference between most of the 'parental' sequences.

In conclusion, a consensus protein is developed using many protein sequences and it is impossible to determine the contribution or value of any of the original 153 input sequences to the eventual consensus design from which specific sequence(s) a consensus sequence has been derived.

(3) Develop the consensus phytase into a commercial product

Next the consensus sequence is back-translated into a corresponding DNA gene sequence which is subsequently synthesised by chemical means or by a combination of chemical and enzymatic means. In general, the codon usage is optimised for the gene to allow efficient transcription and translation in the production organism of choice. Ultimately, the gene encoding the consensus phytase is transformed into the production organism. For phytase, usually a fungal production organism is preferred, but yeasts can also be used.

The genetically modified production organism is then grown in a fermenter, to produce the consensus phytase. Proteomics studies as well as activity assays are used to analyse whether the consensus phytase is expressed, active, and indeed as thermostable as desired.

INTERNATIONAL CHAMBER OF COMMERCE

The world business organization

The International Chamber of Commerce (ICC)

ICC is the world business organization, whose mission is to promote open trade and investment and help business meet the challenges and opportunities of an increasingly integrated world economy.

With interests spanning every sector of private enterprise, ICC's global network comprises over 6 million companies, chambers of commerce and business associations in more than 130 countries. ICC members work through national committees in their countries to address business concerns and convey ICC views to their respective governments.

ICC conveys international business views and priorities through active engagement with the United Nations, the World Trade Organization, the G20 and other intergovernmental forums.

Close to 3,000 experts drawn from ICC member companies feed their knowledge and experience into crafting the ICC stance on specific business issues.

www.iccwbo.org

INTERNATIONAL CHAMBER OF COMMERCE

33-43 avenue du Président Wilson, 75116 Paris, France

T +33 (0)1 49 53 28 28 **F** +33 (0)1 49 53 28 59

E icc@iccwbo.org www.iccwbo.org