United States Department of State
Bureau of Oceans and International
Environmental and Scientific Affairs
Washington, D.C. 20520

1 December 2017

Dr. Cristiana Paşca Palmer
Executive Secretary
Convention on Biological Diversity
413 Saint-Jacques Street, Suite 800
Montréal, Québec, Canada H2Y 1N9

Dear Dr. Paşca Palmer:

The United States appreciates the invitation to provide comments in response to the Secretariat's 8 November 2017 Notification No. 2017-115 regarding the Peer Review of the Fact-Finding and Scoping Study on Digital Sequence Information on Genetic Resources.

We are pleased to have the opportunity to share our views attached.

Thank you for your consideration.

Sincerely,

Barbara M DeRosa Joynt

Barbara M. De Rosa-Joynt
U.S. National Focal Point for the
  Convention on Biological Diversity
U.S. Department of State
Washington, DC 20520 U.S.A.

Attachment: U.S. Submission on Peer Review of Fact-Finding and Scoping Study on Digital Sequence Information on Genetic Resources

TEMPLATE FOR COMMENTS

Contact information	
Surname:	Aktipis
Given Name:	Stephanie
Government (if applicable):	United States Government
Organization:	U.S. Department of State
E-mail:	AktipisS@state.gov
Title of document reviewed:	The Emergence and Growth of Digital Sequence Information in Research and Development: Implications for the Conservation and Sustainable Use of Biodiversity, and Fair and Equitable Benefit-Sharing – A Fact-Finding and Scoping Study Undertaken for the Secretariat of the Convention on Biological Diversity

Comments on the draft fact-finding and scoping study

Page #	Line #	Comment
0	0	This is an example of an entry of a general comment
3	6	This is an example of a specific comment on Page 3, line 6
8	16-19	While the WHO PIP Framework uses the phrase “genetic sequence data,” it does not specifically define that phrase; rather, Section 4.2 defines “genetic sequences”. This text should therefore be changed to the following: “The WHO PIP Framework uses the term *genetic sequence data*, and defines the term *genetic sequences* to be: “The order of nucleotides found in a molecule of DNA or RNA ...contain[ing] the genetic information that determines the biological characteristics of an organism or a virus.”
8	24	The reference to “natural or synthetic” is unclear in that these descriptors do not apply to information but rather to the physical substrates that the information describes or represents. Recommend replacing ‘It may be natural or synthetic’ with ‘It may characterize DNA or RNA that is found in nature, that is designed, mutated, or degenerated, or that is purely hypothetical.’
8	27	Recommend deletion of ‘conventional’. The text appears to be differentiating computational work from laboratory work of any kind, but it is not clear what a ‘conventional’ laboratory is.
9	15-20	The term “synthetic biology” is introduced here but not defined. Please reference the definition in the COP13 decision 17 where “Parties *consider* it useful as a starting point…” This definition should be adhered to throughout the document to avoid confounding the broad field of synthetic biology with genome synthesis techniques or applications.
9	17-20	Recommend inserting “for use in applications” before “such as”. The list after “such as is a list of applications” is not a list of the specific bio-based products that might ensue from synthetic biology. (For example, synthetic biology is not likely to constitute electronics, although synthetic biology-produced materials or organisms may be used in electronics.)
9	23	Clarify what is meant by “reach significance faster”.
10	20	After “remain confidential,” please insert the sentence “Patent Offices also submit sequences included within patent application publication and patent disclosures to databases.” For example, the U.S. Patent and Trademark Office submits sequences (proteins, RNA, and DNA) to GenBank.

10	20	The World Intellectual Property Organization has standards for the presentation of nucleotide and amino acid sequence listings in patent applications, ST. 25 and ST.26. See http://www.wipo.int/export/sites/www/standards/en/pdf/03-25-01.pdf and http://www.wipo.int/export/sites/www/standards/en/pdf/03-26-01.pdf Recommend inserting this as a footnote to the sentence we recommend above (for page 10, line 20).
10	35-36	Given the amount of research that is still based on material samples, “most” over-states the amount of research done using only genetic sequence data. We therefore recommend changing “most research” to “some research” as shown below. “Most research is based on sequences accessed through databases or parts registries, but some groups sequence and analyze physical samples…” to “Some research is based on sequences accessed through databases or parts registries, with many groups sequencing and analysing physical samples …” Also please add at the end “Sequencing genetic materials from organisms being studied in the laboratory is a standard research technique.” It is important to capture the generic and ubiquitous nature of digital sequence information. This is not something restricted to field prospecting or synthetic organism creation.
11	18	The inability of digital sequence information to substitute for actual laboratory manipulation of organisms needs to be emphasized. We recommend replacing “provide” with “are essential for providing”.
11	20	Recommend replacing ‘genome’ with ‘genome sequence’
11	21-32	This example of just one “recent technological advance” is lacking context. The technology outlined in this section, “Oxford Nanopore’s MinION,” is not readily/commonly used or available to researchers working with genetic/genomic resources. Though it is theoretically possible to generate “digital sequence information” in the field, the cost associated with preparing the samples for sequencing and consumables associated with these products make it unlikely that this technology in its current state will be accessible to the average researcher. If references to MinION are retained, we recommend that further context be provided regarding why MinION was specifically identified as well as descriptions of the frequency of its use and its reliability. It is also important to make clear that in order to generate quality sequences, you need to have appropriate physical material with high quality DNA (i.e., not too degraded). Further, the “digital to biological converter” (line 27) is a new technology just announced this year – it is only proof of concept. We recommend including the reference to the Nature Biotech paper describing this technology and provide caveat that it is experimental at the moment. We recommend deletion of the last sentence in this paragraph (“it will soon be possible for most researchers to inexpensively synthesize DNA in their laboratory”). It is misleading in two ways – it is already possible for researchers to inexpensively synthesize DNA in their laboratory for relatively short-stranded oligonucleotides. However, although there are some desktop synthesizers available, the widespread ability to synthesize high-quality gene-length DNA (thousands of base pairs) is not necessarily imminent.
11	29-30	This sentence is vague; please clarify the company or products to which are being referred with the term “synthesizers”.
11	35	This text fails to capture the vast “–omics” libraries, already described, that make their information freely available without restriction. The GenBank website, for example, says that “NCBI places no restrictions on the use or distribution of the GenBank data.” The impression given by this paragraph is that there is a thicket of restrictions. Therefore, suggest keying to line 35 the following: After “…the use of digital sequence information” please insert “, although the major repositories of genetic information such as GenBank provide it for free to all without restriction.”

12	15	It is a misconception that open source materials are not subject to intellectual property rights. Generally, it is the copyright system that makes open source licenses work. It is suggested that "but intellectual property rights will not be asserted against the materials if the conditions of the open source license are met" be used to replace "but materials are immune from the assertion of intellectual property".
12	1	On page 12, line 1, "that might include" is unclear. On page 10, lines 23 and 24 it is said that meta data is rarely contained that includes information on origins of sequences, thus the country of collection would appear to be often unknown. How would the assertion then work?
12	35	Recommend that "barcode" be placed in quotes since "DNA barcode" is not a standard term.
13	4	Recommend rephrasing "understanding pollinators" to something more specific
13	29-38	Recommend deleting this paragraph addresses issues outside the scope of this report and the DSI discussion.
14	2-5	As there is not a system overseeing the management, dissemination and use of DSI for monetary benefit sharing, we recommend deletion of the clause "the system that manages, disseminations, and uses". If the sentence is intended to describe a future system, we recommend changing "the system that manages, disseminates, and uses…" with "a system that were to manage, disseminate, and use …"
15	15	In order to clarify concerns about this issue, after "might isolate data or reduce the effectiveness of databases," please insert ", such as by impeding utilization of databases for research that would never be expected to lead, except perhaps serendipitously, to the development of lucrative commercial products."
15	16	Recommend replacing "many" with "some".
15/16; 16/17	Lines 37-10 and 40 - 6	The paragraph bridging pages 15/16 discusses public databases, indicating that they - "do not require contributors or users to register or log in, agree to terms and conditions, or sign user agreements". The paragraph bridging pages 16/17 states that "sequences move fluidly between commercial and non-commercial institutions, and once uploaded to public databases are available for all to use". The *quid pro quo* of the patent system is that a limited monopoly is provided in exchange for making public the means to make and use an invention. Public availability of patent sequence information is a necessary part of this patent system fundamental. The United States Patent and Trademark Office (USPTO) sends "Pre-Grant Published" and "Grant" sequence listing data (including DNA, RNA, and amino acid sequences) to NIH weekly for posting on the NCBI website. These sequences are available at this website: https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome The search set can be chosen from a number of databases, including "Patent sequences (pat)". While some sequences from published patent applications or granted patents may be in the public domain, other sequences may not. Therefore, simply because certain sequence information may be accessible and searchable on a public database does not mean that use of the information is free from restrictions. In addition, the European Patent Organization (EPO), the Japan Patent Office (JPO), and the Korean Intellectual Property Office (KIPO) also export "Pre-Grant Published" and "Grant" sequence listing data to public databases, i.e. EMBL-EBI and DDBJ. Page 19, 3rd paragraph, notes that "the large databases joined into the International Nucleotide Sequence Database Collection consortium" (INSDC).
16		Recommend including a reference to depositories of microorganisms such as those used by the patent system (Budapest depositories such as ATCC: https://www.atcc.org/) or seed collections such as NCGRP: https://www.ars.usda.gov/plains-area/fort-collins-co/center-for-agricultural-resources-research/plant-and-animal-genetic-resources-preservation/docs/seed-collections/?

16	27	“Watermarking” is the construction of a physical segment of DNA that contains a specified set of base pairs. It may be used to label a physical strand of synthetic DNA but, by its very insertion, transforms naturally occurring DNA to laboratory-modified DNA. Moreover, it says nothing about who may have had access to the sequence information since it was first generated. Replace “watermarking a DNA sequence in a non-coding region of DNA” with “watermarking a physical sample of DNA in a non-coding region of DNA to create a synthetic piece of DNA whose provenance might be identified if it is later sequenced. By definition, naturally occurring DNA contains no watermarks. Moreover, such a watermark says nothing about who may have had access to the sequence information since it was first generated from the physical sample.”
17	20	Recommend deleting “and the future is unclear” as it implies that physical samples may not be needed ever in the future. This seems quite unlikely. It is more appropriate to simply say that their role in research and commercialization is changing.
17	28-29	There already are many benefits from the use of DSI. Recommend saying “that ensure CONTINUED benefits” to make this clear.
19	35-36	The scoping report was commissioned by the ITPGRFA Secretary on the advice of the Co-chairs of the Ad hoc Open-ended Working Group, and not at the request of the Treaty’s Governing Body. We recommend replacing the sentence as follows: “Welch et al (2017) elected to use the term “sequence data” in the scoping report recently commissioned by the ITPGRFA Secretariat.”
20	1-11	While the WHO PIP Framework uses the phrase “genetic sequence data,” it does not specifically define that phrase; rather, Section 4.2 defines “genetic sequences”. This text should therefore be changed to the following: “The WHO PIP Framework uses the term *genetic sequence data*, and defines the term *genetic sequences* to be: “The order of nucleotides found in a molecule of DNA or RNA ... contain[ing] the genetic information that determines the biological characteristics of an organism or a virus.”
20-21	Table 1	This table is confusing taken out of the context of the original paper. Specifically, including “genetic sequence data” as an example of “raw data” in the “explanation” column and also including it in the “types of data” column is confusing, as is repeating the “automatic gene/transcriptome annotations” and “protein sequence data” in the “data only” and “data and analysis” sections. We suspect the original paper explained why the issues were repeated, but without the necessary explanation, this is unclear and should be remedied. Moreover, “protein sequence data” or “protein structure data” does not necessarily constitute annotation with a putative gene function, since the function of the protein may not be known.
21	Table	Why note “embargo” after “Protein structure data?
21	15-20.	As the examples from the Peruvian Society of Environmental Law do not mention digital sequence data, recommend deleting this section as these examples are confusing in the context of a paper on genetic sequence data.
22	1-2	DNA as a storage medium for information is theoretically possible but is orders of magnitude away (in cost and storage/retrieval time) from constituting a practical information storage mechanism. We therefore recommend changing “are now stored” with “can now be stored” and inserting at the end “However, such approaches are orders of magnitude away from being competitive with existing information storage technologies.”
22	17	Please explain what genetic “parts” are as this is a new term for this paper.
22	20-21	Suggest deleting “Earlier discussions focused on DNA sequences, but today sequence information is generally considered to extend beyond DNA.” All the sequence discussions in this paper refer either to DNA sequences, or to protein and other sequences that are derived directly from DNA sequences and can be fully described in terms of their underlying DNA. Extending “sequence information” to more qualitative descriptions and annotations risks capturing other sources of information such as published scientific papers and textbook that go well past what can be derived directly from the genome.
22	33-35	Recommend deleting this sentence as unpublished and “forthcoming” papers should not be included in this report.
22	38-40	The definition of genetic resources in the CBD is not accurate and should be corrected.

23	40-42	The reference to "natural or synthetic" is unclear in that these descriptors do not apply to information but rather to the physical substrates that the information describes or represents. Recommend replacing 'It may be natural or synthetic' with 'It may characterize DNA or RNA that is found in nature, that is designed, mutated, or degenerated, or that is purely hypothetical.'
25-26	22 – 45; 1-30	This section describes synthetic biology and how integral gene sequences are to it. However, some of the examples are overly specific and tend to reduce broad, generalizable and evolving technologies to oddities (i.e., yellow glowing silk). Moreover, synthetic biology should not be discussed independently of its objectives or products, such as agricultural biotechnology, industrial biotechnology, or healthcare. This discussion need not be repeated in all three sections and we recommend an alternative organization, specifically, replacing the heading "synthetic biology research" with the term "synthetic biology technologies" and including the text from 3.1, 3.1.1, and 3.1.5 in this section. We also recommend adding another section called "synthetic biology applications" which would include sections 3.1.2, 3.1.3, and 3.1.4 We also note that an in depth consideration of synthetic biology research and applications is the ongoing work of the AHTEG on synthetic biology and the current DSI study should not pre-empt or bias any of those deliberations or conclusions. We also note that this discussion of synthetic biology does not adhere to the CBD Synthetic Biology AHTEG's definition of synthetic biology. We recommend that all discussion related to synthetic biology remain consistent with the CBD Synthetic AHTEG's definition.
25	39	Using genes from multiple organisms is a possibility but is not inherent to synthetic biology. We recommend changing "taking genes from a number of different organisms" to "taking genes, perhaps from a number of different organisms".
25	40	Replace "an artificial DNA construct – which allows the genes to work together." to "a DNA construct, typically based on a naturally occurring virus that can convey the genes into the organism where they are to be expressed."
25	44-45	We recommend deleting the sentence. "Both the vector and the hosts are also often owned by companies that have associated intellectual property." Hosts are usually E. coli or some other common bacterium. Vectors are a ubiquitous biological tool. What is the relevance of pointing out IP here?
26	1-2	We also recommend deleting this sentence, as it is just one example of what microbes can do: "This process turns microorganism hosts into biological or microbial 'factories' fed by biomass feedstocks that produce sugars." Microbes don't all end up producing sugars or being "factories."
26	27	In line 27, "cost" should be "coast".
26-27	31-41; 1-2	Please make the connection to DSI clearer by providing more information on how DSI contributes to it (right now only enzymes and micro-organisms are highlighted).
27	6-9	The location of healthcare biotechnology is not relevant to this discussion; please delete these sentences.
27	10	Digital operations such as the ones described are certainly becoming more important to life science companies. However, saying that they are "focusing their strategies" on digital minimizes the continuing necessity of laboratory and clinical investigations that test the properties of drug candidates and biologics. There are some applications that are primarily digital – say certain approaches to precision medicine, which perhaps can be based solely on data analytics. But for the most part, information analysis helps develop chemical or biological substances that still have to be evaluated the "old fashioned" way. We therefore recommend replacing "focus their strategies on" with "utilize"
29	23	We recommend including the number of participants (3,000) and the number of countries they represent, as well as the number of teams (310 student-led teams from 44 countries), who participated in the 2017 iGEM competition in the 2017.
29	26-27	Patents and patent application publications are also sources of genomic information, and may also have supplementary files associated with that information. See: http://seqdata.uspto.gov/
31	11	This paragraph states that "All INSDC partners are publicly funded by their host governments, and the INSCD's policy…emphasizes the mandate to free, unrestricted access to all of the data records in their database (Cochrane et al, 2016)." It is important to note, however, that while **access** to all the data records in their databases is free and unrestricted, such free access does not directly translate to **unrestricted use** of patent sequence data. We recommend that this point be clarified throughout this document.

31	33	After "compares nucleotide protein sequences to sequence databases" insert "to identify similar sequences already in the database and".
31	38-41	We note that the problem mentioned here – that a database of sequences might contain identical sequences from different sources, which would then complicate an ABS system – is not a fault of BLAST but rather a fundamental characteristic of life on earth. Genetic functions are not uniquely attached to geographic locations on the earth; they are attributes of living organisms with sometimes extensive geographic ranges, and they may share those genetic sequences with other organisms that are found in other locations. This is a fundamental complication in any approach to providing access and benefit sharing to genetic resources when that term is defined to include genetic sequence information. We also note that this point may be better inserted in the geographical discussion at the end of page 33; similar points are also made on p. 54, lines 7-9, and elaborated on at p. 56, line 10 through p. 57, line 7, but they should be made here as well. For this location, we recommend deleting the last sentence of the last paragraph on this page ("But BLAST might also help researchers find an identical sequence in a different organism as a way of avoiding the use of a sequence that raised legal uncertainties under ABS, or to avoid monitoring (Welch 41 et al, 2017; Bagley, 2017)") and adding a new paragraph, as follows: "Genetic databases may contain sequences from organisms that can have extensive geographic ranges, and identical sequences might be found in different organisms found in still other locations. This fact is an inherent complication in any ABS scheme, since there may be no way to attribute a genetic function to a location of origin."
32	26-27	This sentence overstates the situation, especially because the next sentence (page 32, lines 27-28 and page 33, line 1) notes that the extent of what is uploaded is difficult to estimate. Please change "many" to "some" and "are not" to "may not be".
34	4-10	Is there a source for this quote? Recommend including a citation.
36	4-13	This example of just one "recent technological advance" is lacking context. The technology outlined in this section, "Oxford Nanopore's MinION," is not readily/commonly used or available to researchers working with genetic/genomic resources. Though it is theoretically possible to generate "digital sequence information" in the field, the cost associated with preparing the samples for sequencing and consumables associated with these products make it unlikely that this technology in its current state will be accessible to the average researcher. If references to MinION are retained, we recommend that further context be provided regarding why MinION was specifically identified as well as descriptions of the frequency of its use and its reliability. It is also important to make clear that in order to generate quality sequences, you need to have appropriate physical material with high quality DNA (i.e., not too degraded). Further, the "digital to biological converter" (line 27) is a new technology just announced this year – it is only proof of concept. We recommend including the reference to the Nature Biotech paper describing this technology and provide caveat that it is experimental at the moment. We recommend deletion of the last sentence in this paragraph ("it will soon be possible for most researchers to inexpensively synthesize DNA in their laboratory"). It is misleading in two ways – it is already possible for researchers to inexpensively synthesize DNA in their laboratory for relatively short-stranded oligonucleotides. However, although there are some desktop synthesizers available, the widespread ability to synthesize high-quality gene-length DNA (thousands of base pairs) is not necessarily imminent.
36	36	"in their lab" would be clearer than "on their lab"
37	11-14	This quote is confusing. What are leads? What was derived? Please clarify. We recommend that you also provide a citation for this and all quotes.

38	Table 2	Again, this table is confusing out of its original context. Please clarify further or delete. The row distinguishing between “small” and “big” science is particularly confusing. At a minimum, please clarify what is meant by these terms.
40	15-39	This section appears to suggest that open source does not depend upon intellectual property. Open source generally does depend upon IP (especially copyright) to make open source licensing agreements enforceable. See: [https://www.copyright.gov/about/office-register/wjlta033117.pdf at 136](https://www.copyright.gov/about/office-register/wjlta033117.pdf). Please clarify, and we also recommend deleting “without overturning existing intellectual property laws on lines 32 and 33. Additionally, we recommend balancing the text addressing IP, as it appears biased. It does not acknowledge that without IP, investors may be unable to provide the funding to scale up investment in order to translate promising discoveries to on the shelf products. Lines 30-31 should be deleted, or balancing language added to communicate that the same concerns with regard to patents causing transaction costs also exists regarding requirements for Prior Informed Consent and Mutually Agreed Terms.
41	13-22	In addition to these examples that are labelled “open source” may want to consider adding other examples that promote access such as WIPO Re:Search: http://www.wipo.int/research/en/
42	4-17	We recommend editing items (5) and (6) about GISAID user access agreement. On (5), there is not a requirement to analyze findings jointly but to make best efforts to involve the originating lab in analyses. (See agreement terms: “**Collaboration.** You agree to make best efforts to collaborate with representatives of the Originating Laboratory responsible for obtaining the specimen(s) and involve them in such analyses and further research using such Data”) We therefore suggest combining (4) and (5) in this way: “(4) make best efforts to collaborate with the originating laboratory and involve them in analyses”. We recommend deleting the text in (6) entirely, as the user access agreement does not include a provision that maintains “common access to the technology derived from the data.”
43	15	Recommend replacing “microbes and invertebrates” with “Bacteria, Archaea, and invertebrates”.
43	17	The accuracy of bar-coding to identify species is questionable and depends heavily on the coverage of the database being queried. Replace “as an accurate means to identify species” with “as a means to identify species”
45	5-6	“Understanding pollinators” is vague and should be replaced with something more specific
45	11-12	“Research underway [….] on bees and their pathogens”. What is this research trying to accomplish?
46	29-37	We note that other examples could be added to this section, such as the use of genetic information from pathogens in food-borne outbreaks to help track down the source of the contaminated food. But pathogen sequence information is also important in medical and pharmaceutical research, including disease surveillance and the development of diagnostics, therapeutics, and vaccines.
46	32-34	The existing text could be read to mean that sequence information from the Zika and Ebola epidemics is used in tracing the origin of pathogens in other epidemics; we think what meant was that it was useful in tracing the origin of pathogens during those (Zika and Ebola) epidemics. Suggest reordering (and slightly rewording) the sentence to read “Tracing the origin of pathogens in emergency situations often draws on sequence information; such approaches were used in responding to the Zika and Ebola outbreaks (Tyler, 2017)” Additionally, please add the reference for Tyler, 2017 in the bibliography.
47	19-24	The sentence on lines 19-20 should acknowledge that biotechnology is already increasing farm productivity and reducing environmental impacts. Recommended replacing “might also increase” with “are” and “reduce” with “reducing”. See this report for further information: 2016. Genetically Engineered Crops: Experiences and Prospects. National Academy of Sciences, Engineering, and Medicine.
47	23	As drafted, this sentence is speculative. Please change “could be used to recreate extinct species” to “might be used to recreate extinct species”.

47	25-40	These concerns are speculative and somewhat removed from the issue of DSI. Please provide more substantiated examples, further clarify the connections to DSI or delete the examples.
48	1-9	If these examples are retained (see comment for page 47, lines 25-40), please delete "might become invasive species, or toxic to other non-target organisms" and replace it with the "may be harmful to non-target organisms or impact genetic diversity". All the harms alleged here are speculative, as is any environmental release, but becoming an invasive species seems more speculative than the others. As a general rule one could expect synthetic organisms to be less fit than native organisms to survive outside the particular environmental niches for which they were designed. Please also delete the final sentence of this paragraph, as many of the risks identified are highly speculative, therefore drawing any conclusion from them and labelling it "fundamental" is inappropriate and misleading. Moreover, many of the potential risks associated with technologies that make use of DSI are in fact predictable and risks can be appropriately managed and mitigated.
50	11-37	The 2nd paragraph states that " Slobodian et al (2017) distinguish between a **public domain approach – eg the INSDC databases** – and an open source approach – eg BioBricks and BiOS - that can include conditions (eg materials must be available for multiple generations of users) while still permitting patenting and commercialization." (Emphasis added). This distinction is misleading as NCBI (an INSDC database) contains patent sequence data, some of which may **not** be in the public domain.
50	23-37	Please provide additional examples and more citations for this section.
50	30	Suggest changing "biosecurity concerns would dictate" to "biosecurity concerns would suggest," as such requirement does not currently appear to exist in GenBank.
50	34-35	It is not clear what "accessed from biological diversity" means in the phrase "Although the sequence data of human viruses is not accessed from biological diversity in the same way as that associated with…." Please clarify.
50	41-42	We recommend ending this sentence after "within the community". The international community has agreed that processes can be patented (note, for example, Article 27 of the TRIPS Agreement). As currently drafted, the authors' attempt to distinguish between processes and products in this sentence is subjective. Also, note the comments regarding open source in relation to page 12 and 40.
51	1-6	Paragraph takes only one viewpoint regarding IP. We recommend that it either be better balanced by also including views of patent holders or deleted entirely.
53	7-13	The quote is unclear. How is a hundred year's worth of work proportional?
53	21	"Many" seems high when the only citations are two papers by the same author. Please replace with "some".
55	26	Check the citation. The name might be Duffield, rather than Dutfield.
57	3-5	Recommend deleting the text "including countries that are not Parties to the CBD," as it does not add anything to the overall point of the sentence which is that many homologous or conserved sequences are found in different regions or countries.
58	12-20	Perhaps this paragraph could be clarified with regard to the patent sequence data on public databases. This paragraph is also unclear because many of these databases do apply conditions to use.
59	42	There is a difference between public domain and public access that is overlooked. We recommend that "already in the public domain" be changed to "that is already publicly accessible, especially that which is in the public domain."
61	10-28	This section represents inquiries into potential options for tracing GSD as adopted approaches under the PIP-FW. This is inaccurate. Though technical/experts groups have evaluated potential options, WHO Member States have not adopted these findings and an active review of how the PIP-FW might address GSD is underway. Please ensure that this paragraph correctly reflects activity under the PIP-FW. We also recommend deleting the first sentence of this section, specifically line 10 – 11.

62	1	“determination” should be “variety denomination”
62	8	A citation should be provided for the conclusion as to what “most countries” do or the sentence should be reworded.
62	11-13	“The primary data associated with biological sequences that is provided to **national patent offices** is also **not yet** comprehensive, **standardized**, timely and meaningful (Jefferson et al, 2015).” Please note the recommended international standard for the presentation of sequence information in patent applications is in the process of being updated to include INSDC annotations to the extent feasible. This updated standard is WIPO Standard ST.26 and can be accessed here: http://www.wipo.int/export/sites/www/standards/en/pdf/03-26-01.pdf
63	2-5	We recommend adding the words “among some” after “raised concerns” as not all groups share the views identified in this sentence. Further, we recommend that the authors seek and represent additional viewpoints about DivSeek in the entire paragraph (page 62 lines 37-41 and page 63 1-7).
63	21	Recommend deleting “and the future is unclear” as it implies that physical samples may not be needed ever in the future. This seems quite unlikely. It is more appropriate to simply say that their role in research and commercialization is changing.
63	29-30	There already are many benefits from the use of DSI. Recommend saying “that ensure continued benefits” to make this clear.
81	FN15	This footnote references the INSDC partner data submission systems. Currently, patent sequence data contains only sequences with no annotations. Once ST.26 (referenced immediately above) becomes effective (expected Jan 2022), it is hoped that sequence annotations will be included in the sequence data made available in the public databases.

Submit your comments to secretariat@cbd.int or by fax at +1 514 288 6588.