

TEMPLATE FOR COMMENTS

Contact information		
Surname:	Seberg	
Given Name:	Ole	
Government (if applicable):		
Organization:	GGBN	
E-mail:	oles@snm.ku.dk	
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 #	Para #	Comment
0	0	This document simply has not got the role, importance and efforts of Natural History Collections (in the broad sense Museums, gardens etc.) right.
0	0	It fails to grasp the key role of phylogeny for interpretation of nearly all aspects of biodiversity and to acknowledge the immense importance of biodiversity (other than micro-organisms) as a untapped resource of new products and inventions.
0	0	While acknowledging the potential of DNA barcodes the document exaggerate it's (current) importance and it's obvious limitations
0	0	It must be stressed somewhere in this document that the concept species are used as a placeholder for entities that are defined widely different within micro-organisms and most eu-karyotes.
8	39	It is an exaggeration to write "... genomes of species ..."! What comes out of this is a hotchpotch of sequences from known species and sequences that for some reason cannot be assigned to any known taxon.
9-10	28-32	I might be worthwhile to mention/discuss other databases like e.g. UniProt, Swiss-Prot, TrEMBL and Protein Information Resource Protein Data Bank, Ensembl and InterPro. Additionally some journal requires deposition of NGS raw reads.
10	22	"... included information on the environmental context..." Such data are usually called meta-data and the most interesting meta-data in a Nagoya-context are geographical data (e.g., Country, administrative unit, local features or simply GPS-data). Use or refer to the somewhat deviation information here as compared to P. 59, line 1-P. 60, line 5 .
10	23	Instead of writing "...which organisms originate." it would be better to write "... which ex-situ samples originate."
10	35	Given the amount of research that still is based on material samples, "most" overstates the amount of research done using only genetic sequence data. We recommend changing "most research" to "some research".

11	5-11	This is not precise enough. I agree that few companies collect field samples – it is simply not cost-efficient. However, all organisations that rely on access to biodiversity, e.g. wild relatives of crops or animals” plus the worlds +3000 Botanic Gardens, the equally or probably much larger number of Natural History Collections (Museums) etc. (see P. 11, line 12-16) actively engage in general collecting the World’s Biodiversity. Citizen science project are indeed important but often narrow in scope.
11	12-16	Most of this paragraph has nothing to do with DSI except the last 7 words
11	17-18	This statement is simply not correct... Industry may be moving away from it, but science in general is hardly doing so. There is a constant need for samples of taxa none had considered important until it is certainly realised that they might include valuable compounds (e.g. as soon as Thapsigargin became medically interesting the need for sampling – even the genus <i>Thapsia</i> ’s taxonomy increased tremendously!)
11	21-32	Are these technologies commonly used right now? We recognize that they exist, but suspect that they are not yet standard in research. As drafted, these paragraphs make it sound like these technologies are extremely common. We recommend that some language is added describing how frequently they are used and their reliability.
12	1	A more precise wording than “A number...” would be “A small number....”
12	30-31	It goes beyond that! It is an essential tool in helping understand the size of the world’s Biodiversity. Maybe surprising so (and it will probably have to be emphasised elsewhere in this document: We have only scanty knowledge of how many species we share this globe with. According to Mora, C. et al. 2011. How Many Species Are There on Earth and in the Ocean? PLoS Biology: e100127, only 13% (1.44 million of the species on Earth are known, leaving 9.52 million totally unknown). This is a serious problem as they are all potentially important to humankind. It is even worse than indicated as a huge proportion of the known species are only known from one or a few specimens. Very, much so for (P. 12, line 31-34) “Understanding the Earth’s biodiversity and its”
12	35	This bullet-point is an exaggeration better to write: “DNA barcodes, used to identify species;”
12	36	This is an emerging field; few countries on the Earth are near to this goal.
12	36-38	This is not correct! Sequence data are usually integrated in taxonomic work, e.g. delineating species. Most – but admittedly not all taxonomists – consider it bad science to rely on one type of data. This work creates the foundation for most use of DSI that work with more than one species; and even there! Much more important DNA data – increasingly genomic data – are an integrated part of constructing the Tree of Life (phylogeny). Why is the Tree of Life important that important? Because phylogenetic relatedness is important, e.g. for conservation of biodiversity, understanding emerging diseases, searching of biologically active compounds, understanding evolution, including (P. 12, line 39) understanding genetic variation in populations, etc. , etc. None of the bullet-points (incl. P. 13, line 1-6) can be fully understood without! “ <i>Nothing in Biology makes sense except in the light of evolution.</i> ” (Theodosius Dobzhansky; 1973).
13	7-38	Must be adapted such that it is in line with the above!
14	16-17	The loss of control is a strange argument, as the Nagoya-protocol is only interpreted as retro-active in a few places control is already lost of most available sequences.
15	16	Replace “many” with “some”.
15	23	The sentence “Additionally, homologous, or identical, sequences ...” makes no sense. It suffices to say homologous or orthologous. Similarity, incl. identity is measure mathematically whereas homology is a concept. Homologous sequences may be quite different.
16	11-13	This is standard in INSDC.

17	8-30	Need to be re-written due to imprecise information given in this document previously (see above).
17	19- 20	Delete "...and the future is unclear" as it implies that physical samples may not be needed in the future. This seems highly unlikely. It is more appropriate simply to say that their role in research and commercialization is changing. They will most likely remain of considerable – and potential – high importance. In fact very much so!
17	28-29	There already are many benefits from the use of DSI. Recommend writing "... that ensure continued benefits..." to make this clear.
24	30	Change "... which may contain thousands ..." to "... which depending on sample size may contain thousands ... "
24	35-39	This applies to "classic DNA barcodes" the science is moving away from this simple interpretation.
29	32-33	Change "... from field collections involving..." to "...from field collections involving samples of biodiversity from the environment and ex-situ collections."
30	25	The part of the sentence reading "... genetic variation data from all species;" is wrong. We are far from having even just one sequence from all known species.
32	26-28	Change "Database managers and others note that many datasets are not entered into international public..." to "Database managers and others note that some datasets may not entered into international public..." Primarily because of the contents of the next sentence!
34-36-	35 3	This section completely underestimates the enormous effort made by the world's Natural History Collection to sample the 9.52 million unknown species. Many of the species will routinely be sequenced during their taxonomic treatment. It is worth emphasising that not only pro-karyotes and micro-organisms but also many eu-karyotes are of potential importance for humans (see above). Meta-genomic may have caught on in part of the industrialized world but hardly anywhere else – Hence, much of the hype about meta-genomics in this text is, albeit a highly powerful technique, a red herring.
35	11-34	These citizen science projects are fabulous and a good way of raising awareness. The will provided very usefully data – but all citizen science projects have their obvious limits. They are only scratching the surface.
37	25	What are genetic sequence collections in this context?
37-38	37 10	Due to their imperial past Kew Gardens and the Natural History Museum, London, are at the cutting edge of adapting to the Nagoya-protocol. Otherwise what they are doing with their collections including which parts of and to what extent the ex-situ collections are mobilised is pretty standard for that kind of institutions.
43	2-10	Rewrite the opening sentence of this paragraph to "... information creates the background information, viz. the taxonomic and phylogenetic framework, and essential to support biodiversity conservation" and continue "These deepen our knowledge about biodiversity, identify ..." [The 2 nd sentence is too long and winded!]
43	13-14	As indicated above ... we have no firm knowledge about how many species we share this Globe with!
43	15-17	As indicated above, this may be true for a few industrialised countries, but not for the rest of the world. DNA 'barcodes' exist for a fraction of the world's biodiversity – mostly specialized uses, e.g. identification of bush-meat etc. DNA barcodes combined with traditional taxonomy is potentially very strong method hardly used to its full potential. Some research groups are already in the process of supplanting the traditional, few DNA marker approach with a whole genome sequences (mitochondrial DNA and plastid DNA) ditto.
43	30	I would definitively remove the word "... significantly ..." - yes, it has contributed! It works extremely well in the case-studies mentioned (P. 43, line 32-line 38).

43	39-41	Microbial taxonomy has next to no similarity to traditional taxonomy, which make it much easier to employ DNA-based technique.
45	24-30	I know these studies well... How even in species poor ecosystems the major bottle-neck is lack of reference libraries, which may be easily produced in such contexts, but with difficulties elsewhere.
45	33-34	This is only partially correct. For historical reasons ex-situ collected material are considered objects in their own right. However, subsampling of these objects has always been done for other investigations. Accordingly, DNA material taken from a toepad of a bird just end up in another sub-collection (in a freezer or in a cryo-tank) with reference back to the specimen it was taken from. Like a microscopical slide ends elsewhere with a link to the specimen it was taken from - all for practical reasons.
45	34-36	Sequence data are not a golden taxonomic bullet as indicated above it is an extra piece of evidence: Sometimes conclusive; sometimes not.
45	36-39	This is highly doubtful except in special cases – even if sequencing was totally free think about the manpower it would take to do this for 80 million specimens in the Natural History Museum, London.
46	3-8	Makes no sense – rewrite and explain better.
46	12	Rewrite: “...e.g., support the fight against...” Comment: DNA barcodes work excellent is such restricted a universe!
47	35	CO ₂ not CO ²
49	35	A request for authorship may be countered by the Vancouver Declaration.
53	21	“Many” seems an overestimate when the only citation are two papers by the same author. Please replace with “some”.
63	20-21	Delete “...and the future is unclear” as it implies that physical samples may not be needed in the future. This seems highly unlikely. It is more appropriate simply to say that their role in research and commercialization is changing. They will most likely remain of considerable – and potential – high importance. In fact very much so!
63	29-30	There already are many benefits from the use of DSI. Recommend writing “... that ensure continued benefits...” to make this clear.
70	12-15	Reference is wrong! Substitute with: Droege, G., Barker, K., Butler, C., Lyal, C., Seberg, O. 2017. GGBN - Strategies for Standardized Exchange of Genetic Resources on a Global Scale. In: Löhne, C., Zippel, E., Rohkemper, M. and Gardt, S. (eds) Genetische Ressourcen, Gesetze & Gute Praxis: Wege zur Umsetzung des Nagoya-Protokolls in Deutschland. Projektbericht. BfN-Skripten

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