CONVENTION ON BIOLOGICAL DIVERSITY (CBD) NOTIFICATION 2017-115

Peer Review of Fact-finding and Scoping Study on Digital Sequence Information on Genetic Resources

Submission by Australia

NOTE: **All information provided in this response has been drawn from Australian Government agencies.**

Notification 2017-115: Peer Review of Fact-finding and Scoping Study on Digital Sequence Information on Genetic Resources

Australia thanks the Executive Secretary for commissioning a fact-finding and scoping study to clarify terminology and concepts and to assess the extent and the terms and conditions of the use of digital sequence information on genetic resources in the context of the Convention and the Nagoya Protocol.

Australia is pleased to provide the following submission in response to Notification 2017-115.

Feedback and Comments on the draft document		
Page #	**Line #**	**Comments and feedback**
0	0	General Comment: there seems to be little information included from sources/input outside of the US and Europe.
11	21-32	Prior to this section it would be useful for the reader to outline that sequence reading and writing are two rapidly-advancing technologies, and why they are important for synthetic biology, which focusses on use of the DSI. This would provide better context for this section. Also, it would be appropriate to outline the full gamut of recent advances in sequencing technology, rather than only mentioning one (MinION instrment, see comments below).
11	21-25	The section on the MinION nanopore DNA sequencing instrument reads a bit like an advertisement for the instrument. Suggest re-phrasing into more appropriate text, and that details about commercial ventures and commercially available technology are probably more appropriate in an annex. In any case, reviews for this machine are somewhat mixed; it would be worthwhile examining its capabilities more carefully, and it probably is not appropriate for this document to appear to be endorsing one commercially available technology over the many others that are available. 'Affordable'/'Low-cost' it may be but I understand that the maintenance and consumables costs are still very high, making it cost-prohibitive for many users. Generally speaking, there have been many recent advances in DNA reading technology, it is surprising that they are not mentioned here at the same time. Introduction of these technologies in general with respect to how they work is appropriate but reference to companies that provide services or instruments should probably go into an appendix or end note
11	27-29	'A digital-to-biological-converter has been developed to produce functional biologics in an automated fashion from digitally transmitted DNA sequences, in particular DNA templates, RNA molecules, proteins and viral particles'. More information and context about this 'digital-to-biological converter' is required here. What is the machine, why is it better than current synthesis technologies (is it better?), who offers this technology,
11	34 -	Section on tools to manage DSI: there should be more discussion on how wide-spread (or otherwise) these tools are. No such tools exist for access to NCBI for example.
12	20-27	An important example to include here is that of the Yeast 2.0 consortium public domain agreement
17	18-19	Another reference to the MinION 'affordable and portable' sequencer
18	24	Three months or four months? It says four months in the Exec Summary
18	37	Important to know which countries and what types of individuals (number of academics, industry reps, etc.) were included in these interviews to help demonstrate who was involved and where they were from, and how broad the consultation was. This would help

		contextualised which 'researchers' the document refers to later in the document, eg. p.21 line 8
21	7	It might be useful to refer readers to CBD XIII/16 (https://www.cbd.int/doc/decisions/cop-13/cop-13-dec-16-en.pdf). Note also that specific definitions of 'digital', 'sequence', and 'information' are missing from this document. We agree that use of the phrase 'digital' seems tautological with respect to 'sequence information' when genetic sequences are what's under discussion
22	11	See notes for page 21 line 7
22	19	17-20: reference to 'functional unit of heredity' is somewhat redundant/anachronistic in this discussion. In synthetic biology, 'parts' refer to discrete sequences that have known identified and reproducible functions ; that might be encoding a protein or controlling expression of a gene or any number of other functions. Also, it's an assertion that 'genetic parts are of most interest to researchers' – they might be of most interest to the synthetic biology community but there is a much larger community which uses DSI and has varying levels of interest/use for the 'parts' concept
22	34	'expressions of natural information other than nucleic acids and amino acids' – this is a vague and not very useful comment in the absence of sufficient context. Who is Vogel? Where is this context 'forthcoming'?
0	0	Section 3 General Comments Data production in terms of the technology available to produce sequence information is well reviewed in the first part of section three (prior to the 'use' comments), but it is not well explored with respect to the amount and types of sequence data being stored. Moreover, there is not a strong understanding conveyed of the quality of the data being stored, which can be highly variable. This might deserve its own section. The 'Synthetic Biology' section in 3.1.1 (pp25-26, lines 39-45 both pages) is a very narrow view of synthetic biology and quite focussed on one application (microbial cell factories). There are myriad other things that can be done with synthetic biology – biosensors, response systems, molecular modification, gene drives, remediation… the list goes on and on. This section should probably be broken down into various synthetic biology applications, of which 'Industrial Biotechnology' – which is given its own section immediately after – would include applications of microbial cell factories. Industrial biotechnology here appears to be defined as bioprocess and biotransformation, but that's extremely narrow. The section headings and text arrangement needs re-visiting in here. The whole of section 3 should probably be referred to as industrial biotechnology, with sub-sections describing different applications 'microbial cell factories, healthcare (with overlaps in which cell factories produce therapeutics etc.).
27, 28	37-38, 1-11	This is a quite narrow view of how DSI can be used in agriculture. Mining of course occurs, this might be used to directly egnieer traits in crop plants via gene editing or more traditional genetic modification, and the outcomes from these experiments might be used to direct more traditional breeding programs aimed at producing and/or introgressing a desired trait. I would also avoid introducing researcher opinion as fact without reference to published research (e.g. 'plant systems for bioproduction, which could be far more productive than 8 microorganisms' – while it may be true it's probably only the case in specific instances)
28-29	2-	This is very focussed on the 'open science' topic and there is not much information on community labs and DIYBio. More information on the latter two, as well as an idea of real size and scope of these movements, would be appropriate
0	0	Section 4 General Comments: There is no section on private / government repositories – what are they, who owns them, how many are there, what other information is available?
31	4	It is not clear on how much data is being stored as raw vs processed, and how the raw data is being linked to the processed data.
31	13	It is not clear on how much data is being stored as raw vs processed and how the raw data is being linked to the processed data.

33	8	Considering that the whole section is about "**digital sequence information sharing and compatibility between databases**" it seems strange that TDWG was not mentioned
34	11-28	If the Registry of Standard Parts is mentioned then it should be described how this Registry is developed – basically, from the iGEM competition (which isn't mentioned anywhere and should be). Also, the basic design rules should be explained so the reader has some idea of what a part is and how they work for construction. The Registry is comprised of components that are developed and tested through undergraduate teams which enter a competition. To my knowledge parts from this Registry are rarely used outside of the iGEM competition – I don't think anyone actually uses BioBricks very much outside iGEM.
0	0	Section 5 General comments Examples used are quite specific in Field Collections. Austrlaian examples include Environomics https://www.csiro.au/en/Research/Collections/Environomics
36	4-13	The MinION gets another mention here, and again it looks like a sales pitch. As noted, there are mixed reviews from users of this instrument, which suggest that these sections should be toned down (I just found out there's one in my institute that sitting in the back of a drawer unused). As noted above, equal attention should be paid to other sequencing advances.
36	18-19	The companies listed are all North American. A bit more context outside this arena would be appropriate (e.g. GeneArt, Genscript, etc.). China is playing an enormous role in DNA synthesis
36	25-28	If Yeast 2.0 is mentioned here in genome synthesis it would be appropriate to mention other synthetic genome projects (prokaryotic) that have already been completed (thus demonstrating proof of principle). The key thing about the yeast project is that it will be the first eukaryotic genome synthesised. The adenovirus mentioned was not the first virus genome produced, and there are several examples of bottom-up genome synthesis, notable from the JCVI.
0	0	Section 6 General Comments This section appears to cover the whole gambit of ensuring the user is aware of the terms of usage of the data, etc. There are many tools and methods to achieve this, they have broadly mentioned a few Missing is information in this section on agreements with traditional land owners and limitation of collection/usage of the data based on their beliefs (might be in Section 8). Also missing is a more fulsome description of standard institutional Materials Transfer Agreements (commonly with reach-through IP clauses etc) and a more extensive discussion of the enormous transaction cost of negotiation and implementation of these MTAs, which is what has driven the lower transaction cost and open sharing platforms
3839	15-21, 1-22	There are very many collections of microbes, it seems a bit random to single out the WFCC and spend such a large amount of text describing their operation. What about CGSC, ATCC, etc.? This section should probably be more generic
40	15	OpenPlant is another champion of open source and diminished transaction costs for sharing information https://www.openplant.org/ it should probably be discussed here also. Note that traditionally open source has worked in the software world, but has also been open to abuse and legally is still a mine field. Replicating this paradigm in the world of managing digital sequence information could be risky. It needs to be clear that the tools to manage the information are separate from the information itself. Open source tools are ok, a version of open source data (i.e. a system that does away with MTA's) may be problematic. It is likely that different sharing models will work for different environments, different sequence collections, different applications, etc.

0	0	Section 7 General Comments: n/a
43	13-29	It is very surprising that the major whole genome sequencing initiatives (vertebrates in genome 10K, insects in i5K, birds, etc) have not been mentioned here. Also, no mention of the value of transcriptome- or exon capture-based methods, nor of whole mitochondrion sequencing, as reduced representation methods that provide taxonomically relevant data for a much lower cost.
44	17-22	Perhaps more important to mention that sequence information can be used to monitor the success of such genetic interventions (Weeks et al. 2017, Nature Communications 8: 1071)
44	40-42	I note the lack of a citation here, which is likely because researchers cannot reliably and effectively predict likely establishment simply from knowledge of source populations. The focus of this statement should be on understanding (and blocking) the introduction pathways, which can be identified from the genomic sequence data.
45	11-12	This section does not really say how genetic sequence data is being used to understand anything about pollinators. For example transcriptomics and proteomics has been used to understand how bees respond to the factors believed to cause honey bee decline. Pollen genetic sequencing is also being used to understand pollination patterns of various pollinator species (reviewed in Bell et al. 2016, Genome 59: 629-640)
47	15	It is probably better to use "genetic sequence information" instead of "DNA sequencing" here, as much of the relevant information/value in this example will have come from expression analyses.
47	22	Perhaps include reversing genetic erosion in cultivated animals and plants, since documenting this issue with GSI was described in a previous section.
47	23	Genome sequence information is also vital to understand how SynBio applications like gene drives are likely to spread, including devising strategies to limit their spread or to confine their effects to particular populations.
47	26-40	This whole paragraph may be appropriate to the position paper on SynBio, but is nothing more than editorialising here - as it has nothing to do with DSI.
48	6-9	This statement would be more meaningful as a call for the need for more research on ecological risks and socioeconomic impacts before technologies are deployed (on which SynBio initiatives have focused much more substantially since the 2015 reference provided here).

0	0	Section 8 General Comments Worth noting here is that it's unclear how well known the Nagoya Protocol is in the broader research community. I commonly come across people who have never heard of it. That definitely constitutes a challenge for fair and equitable benefit sharing, the subject of this chapter It would be of value to introduce more specifically the aims and framework of the Nagoya Protocol at the beginning of this chapter, otherwise a lot of assumptions are being made. This section is well written and covers a very broad gamut of issues comprehensively
48	13	This section suddenly starts using the 'ABS' acronym a lot – a reminder of what it stands for is probably useful here.
48	33-39	Australia is not mentioned (were any Australians interviewed for this document? If not it is probably an oversight and you are missing input from a whole continent!). Possibly Australia deserves a mention here because of a specific focus on biodiversity/conservation applications of SynBio (https://research.csiro.au/synthetic-biology-fsp/).
63	19	Yet another reference to the MinION instrument. Please see comments above regarding this