MEETING OF THE STATES PARTIES TO THE CONVENTION ON THE PROHIBITION OF THE DEVELOPMENT, PRODUCTION AND STOCKPILING OF BACTERIOLOGICAL (BIOLOGICAL) AND TOXIN WEAPONS AND ON THEIR DESTRUCTION

BWC/MSP/2008/INF.1
28 November 2008

ENGLISH ONLY

2008 Meeting
Geneva, 1-5 December 2008
Item 7 of the provisional agenda
Consideration of oversight, education, awareness raising, and adoption and/or development of codes of conduct with the aim of preventing misuse in the context of advances in bio-science and bio-technology research with the potential of use for purposes prohibited by the Convention

BACKGROUND INFORMATION ON SCIENTIFIC AND TECHNOLOGICAL DEVELOPMENTS THAT MAY BE RELEVANT TO THE CONVENTION

Submitted by the Implementation Support Unit*

Summary

This document summarises scientific and technological developments potentially relevant to the Convention that have come to the attention of the Implementation Support Unit in the course of its research on the oversight of science, in preparation for the 2008 Meeting of Experts and Meeting of States Parties. Developments covered in this document include genomics technologies, synthetic biology and the open-source publication of raw research data.

I. Introduction

1. Reviews of relevant scientific and technological developments are carried out every five years at review conferences. Information is provided on an ad hoc basis by States Parties. The last review was carried out at the Sixth Review Conference in 2006. Ten States Parties contributed papers on this issue in 2006[1]. Fewer states contributed to earlier reviews.

* Submitted after due date, as soon as required information was available to the Secretariat for inclusion.

[1] The ten countries were: Argentina , Australia, China, Czech Republic, Netherlands, Portugal, Russian Federation, Sweden, United Kingdom, and United States of America. The papers can be found on the BWC website at: http://www.unog.ch/unog/website/disarmament.nsf/(httpPages)/23958FD3E9A0A67BC12571F10032D47B?OpenDocument&unid=3496CA1347FBF664C125718600364331

2. Major relevant scientific and technological developments, identified in 2006, were summarised in a background information document produced by the Secretariat[2]. That document discussed elements of: biotechnology, such as bioprospecting, high-throughout screening, and biological microprocessing; genomics, such as DNA sequencing, DNA synthesis, DNA silencing, DNA shuffling and genomic medicine; proteomics, such as high-affinity binding reagents; bioinformatics and computational biology; systems biology, such as bioregulation; drug discovery, design and delivery, such as combinatorial biochemistry, rational drug design, drug targeting, microencapsulation, biopharming and bioproduction; synthetic biology and biological engineering; as well as other developments, such as nanotechnology, gene therapy, genetic engineering of viruses, anti-viral drugs, detection technologies and biological pest control.

3. The next comprehensive review will take place at the Seventh Review Conference in 2011. During the course of its activities in 2008 the ISU identified three areas of scientific progress which may be of interest to the 2011 review: synthetic biology; gene silencing; and open-source publication of raw research data.

II. Genomic Technologies

4. Of the genomic technologies identified in the review of relevant science and technology in 2006, three areas have developed significantly over the intervening years: gene sequencing, gene synthesis and gene silencing. A useful review of various gene technologies was published by Nature Publishing in October 2007[3].

Gene Sequencing

5. As noted in 2006, "sequencing is the identification of the order of the nucleotides which make up genetic information"[4]. The order of these nucleotides determines the heritable genetic characteristics of an organism – its genes. These genes in turn control many of the characteristics of what an organism looks like and how it functions. Mapping the sequences of an organism, therefore, allows all its constituent genes to be identified. This catalogue of genetic information is known as an organism's genome.

6. Commercial applications for gene sequencing technologies are also beginning to appear. Being able to sequence an individual's genome would allow a review of their biological functionality, the identification of pre-disposition to certain medical conditions (therefore enabling increased surveillance for those conditions, allowing for early identification and initiation of treatment – often a critical factor for successful treatment), and the development of personalised medicine through pharmacogenomics (such as creation of drugs for an individual's specific biology – increasing efficacy and decreasing side-effects). Such developments rely on being able to read genetic information cheaply. The costs involved often hinge upon the time

[2] Background Information Document on New Scientific and Technological Developments Relevant to the Convention, BWC/CONF.VI/INF.4

[3] Nature Milestones, DNA Technologies, October 2007, http://www.nature.com/milestones/miledna/index.html

[4] Background Information Document on New Scientific and Technological Developments Relevant to the Convention, BWC/CONF.VI/INF.4

taken, and the amount of physical intervention needed. The more automated and the faster a process is, the cheaper it becomes.

7. The trend, identified in 2006, towards increasing automation in gene sequencing has continued. Advances in technology continue to increase the throughput of automated gene sequencers. In December 2007, the *Economist* newspaper noted that a singe gene sequencer was capable of sequencing the human genome (about 3 billion nucleotides in length) in two months[5]. A day's output from a first generation sequencer could be replicated, at the end of 2007, in less than 10 seconds. The costs involved have also continued to fall. When the preliminary sequences of the human genome were released in 2000, they had cost millions of dollars. It was reported in March 2008 that a commercial biotechnology company in California, USA had sequenced a human genome for $60,000, excluding labour[6]. The company and scientific commentators alike predict that prices will continue to fall.

8. There are rewards for working on increased automation, accuracy and speed and decreased costs. In addition to the commercial applications, the X Prize foundation, is now offering a $10 million prize for the first team to sequence 100 individual genomes with an accuracy of 99%, within 10 days. Each sequence is to contain at least 98% of the genome and cost $10,000 or less[7]. The competition is due to run until 2013 but the organisers are already considering extending the deadline.

9. While this technology will undoubtedly provide numerous economic, social, health and agricultural benefits, it also has the potential to assist in prohibited activities. A more detailed understanding of how biological organisms function enables better efforts to disrupt that functionality, as well as enhance it.

Gene Synthesis

10. If gene sequencing is "reading" then gene synthesis is "writing". It is the ability to take sequence information (on paper or digital) and create the physical nucleotide sequence that it lists. As with sequencing, developments in automation and throughput have increased the accuracy and length of sequence that gene synthesisers can produce, while reducing the costs and manual labour involved (as well as the levels of technological understanding required to use them).

11. These technologies have already found commercial applications, and for-profit gene synthesis companies have been established throughout the world. These companies allow research groups to acquire gene sequencing capabilities without having to purchase all the necessary equipment. Outsourced gene sequencing plays an increasingly important role in the scientific endeavour but presents unique challenges for regulation.

---

[5] DNA Direct, The Economist, 6 December 2007

[6] Genome is a snip at $60,000, New Scientist, 25 March 2008, http://www.newscientist.com/article.ns?id=mg19726482.900

[7] Inside the Forgotten X Prize – the one that can save your life, Popular Mechanics, 25 April 2008, http://www.popularmechanics.com/science/health_medicine/4260863.html?nav=RSS20

12. Developments in gene sequencing have also enabled the synthesis of *de novo* organisms. This has resulted in the creation of pathogens from nothing more than raw information (i.e. the research group did not posses the pathogen to begin with but only its genomic sequence, which can be obtained for free via the Internet, without even having to register). It has also enabled the creation of custom-made or tailored new organisms (that have not previously existed in nature). Both developments appear to be particularly relevant to the Convention.

13. The first ground-breaking work on the *de novo* creation of an organism was reported in 2002 by Cello *et al* and related to the virus that causes polio[8]. Accomplishing this task took a full laboratory staff two years continuous work. Eighteen months later, a much smaller team at the Venter laboratories managed the *de novo* creation of a bacteriophage virus (a virus that infects bacteria) about three-quarters of the size of the polio virus, in two weeks[9]. In 2005, a different team managed to use these techniques to re-create an extinct virus, the 1918 influenza virus (roughly twice the size of the bacteriophage)[10]. In 2007, the virus responsible for the SARs outbreaks was added to the list of synthesised pathogens[11]. This marked an important step as the coronaviruses involved are over twice the size of the influenza virus and almost four times the size of the polio virus.

14. These experiments, and others like them, demonstrate a real capacity for the *de novo* creation of viruses. This has led to proliferation concerns, especially as it relates to the ability to acquire the most dangerous pathogens. They are no longer confined to nature and a small number of tightly controlled facilities. Questions have been asked about the feasibility of obtaining an agent such as smallpox using these techniques. (Smallpox has been eradicated in nature and stocks are maintained in only two locations.)

15. While larger and more complex viruses are being successfully synthesised by an increasing number of research teams, one group at least has sets its sights at the next threshold - the *de novo* synthesis of a bacterium. Bacteria are several orders of magnitude larger and more complicated than viruses. But levels of complications and size and not all that such an effort will have to confront. The creation of a bacterium in this way may pose additional ethical concerns.

16. During the second half of 2007 and the first months of 2008, research teams at the J. Craig Venter Institute (JCVI) have reported almost all the preparatory steps for the *de novo* synthesis of a bacterium. In June 2007, JCVI applied for a patent on a synthetic minimal bacterial genome[12]. This stripped-down organism contained the "minimal set of protein-coding genes which provide the information required for replication of a free-living organism in a rich

---

[8] For more information on this experiment, see: Background Information Document on New Scientific and Technological Developments Relevant to the Convention, BWC/CONF.VI/INF.4

[9] Smith et al, Generating a Synthetic Genome by Whole Genome Assembly: ψX174 Bacteriophage from Synthetic Oligonucleotides, Proceedings of the National Academies of Science, Vol. 100, 2003.

[10] Tumpey et al, Characterization of the Reconstructed 1918 Spanish Influenza Pandemic Virus, Science, Vol. 310, 2005.

[11] Rockx et al, Synthetic Reconstruction of Zoonotic and Early Human Sever Acute Respiratory Syndrome Coronavirus Isolates that Produce Fatal Disease in Aged Mice, Journal of Virology, Vol.81 No.14, July 2007.

[12] Philipkoski, Scientists Apply for First Patent on Synthetic Life Form, Wired, 7 June 2007, http://blog.wired.com/wiredscience/2007/06/scientists_appl.html

bacterial culture medium"[13]. It is a basic vessel in which a tailored genome can be inserted to replicate. In August 2007, a team at JCVI reported being able to transplant the genome from one species of bacteria to another[14]. The host bacterium becomes the donor species, after amalgamation. In theory, this suggests that it should be possible to insert a suitable, tailored genome into a replication vessel and not only will the replication vessel accept the genome but should adapt to meet the requirements of the new genome. In January 2008, JCVI reported that it had succeeded in the chemical synthesis of a bacterial genome[15]. Although this amounted to creating all the relevant genetic information from scratch, it did not mean creating the whole organism. It remains the largest piece of genetic information assembled to date. The genome was assembled via a number of techniques in other bacteria and yeast and did not result in a functioning organism. In an interview, the research team indicated that it will be necessary "to discover whether cells can be 'booted up' into action when loaded with this genetic programme"[16]. Efforts to combine these steps and to create a bacterium have continued. Representatives from JCVI are expecting to have assembled a fully replicating bacterium 'any day now'. Similar techniques are already being developed for tailor-made or engineered organisms (see Synthetic Biology) and are already being directed towards commercial ends.

Gene Silencing

17. The background paper on relevant scientific and technological developments, in its consideration of gene silencing, outlined the role played by RNA inference (RNAi) as "interfering with the process by which genetic material is read and converted into a product"[17] thereby allowing the function of the genetic material to be established. The importance of these techniques was confirmed in 2007, when the scientists that developed the original "knock-out" technology were awarded the Nobel Prize in Medicine.

18. In a recent review of Advances in Biotechnology and Future Impact on Animal Health, the OIE outlined a range of different molecules used in RNAi. Each of the molecules has different characteristics and structures:

> "As independent small double stranded RNA molecules, short-interfering RNA (siRNA) can only survive transiently in an animal. For prolonged activity the siRNA must be incorporated into an expression vector that functions like a gene. In this case they are called short-hairpin RNA (shRNA) due to their physical structure, or microRNA (miRNA). The latter are found naturally in animals and are now believed to play a major role in

---

[13] Glass et al, Minimal Bacterial Genome, US Patent Application 20070122826, 31 May 2007, http://appft1.uspto.gov/netacgi/nph-Parser?Sect1=PTO1&Sect2=HITOFF&d=PG01&p=1&u=%2Fnetahtml%2FPTO%2Fsrchnum.html&r=1&f=G&l=50&s1=%2220070122826%22.PGNR.&OS=DN/20070122826&RS=DN/20070122826

[14] Lartigue et al, Genome Transplantation in Bacteria: Changing One Species to Another, Science, Vol. 317, 3 August 2007.

[15] Gibson et al, Complete Chemical Synthesis, Assembly and Cloning of a *Mycoplasma genitalium* Genome, Science, Vol. 319, 29 February 2008

[16] Ball, Genome stitched together by hand, News, Nature, 24 January 2008.

[17] Background Information Document on New Scientific and Technological Developments Relevant to the Convention, BWC/CONF.VI/INF.4

normal cell growth and survival. Synthetic miRNA can also be constructed in the lab for delivery to an animal."[18]

19. A review of applications for RNAi in fighting disease, published by Nature News in March 2008 suggests that miRNAs offer opportunities for suppressing entire biological pathways rather than just the products of a single gene – providing "novel therapeutic [avenues for diseases] that are not amenable to other applications"[19]. However, other research reviewed in this article suggests that while RNAi is obviously effective, we may not yet fully understand its mechanism of action.

20. RNAi is already enabling researchers to reduce cholesterol levels in test primates[20] and has been linked to a possible cure for Hepatitis C[21]. Commercial applications for RNAi are already under development. A 2008 review of the development of MicroRNA-based therapies currently under development lists five genomics and therapeutics companies working on treatments for cancer, cardiovascular disease, muscle diseases, viral diseases and for use in diagnosing cancer[22]. The following month, Nature News, reported that RNAi-based therapies had "gone mainstream". In addition to the small biotechnology companies active in the area, one of the large pharmaceutical companies, GlaxoSmithKline, had negotiated a deal to develop MicroRNA-based drugs for inflammatory diseases[23]. The involvement of the large pharmaceutical companies suggests that this technology is coming of age.

III. Synthetic biology

21. Developments in genomic technology, as well as in other disciplines, have enabled a more refined understanding of biological processes and the creation of tools to be able to manipulate them. These advances have prompted some to start to consider developing an engineering approach to biology, an approach dubbed *synthetic biology*. The ability to understand mechanisms and to manipulate them in a systematic manner enabled the establishment of mechanical engineering; similar advances in chemistry led to chemical engineering, and breakthroughs in handling electricity spawned electrical engineering. Now it might be the turn of biology. The discipline of synthetic biology, the existence of which seemed to be disputed in 2006, now seems well established.

22. Current developments are, for the first time, providing opportunities to stop thinking "how can a specific biological resource be of benefit?" and to start thinking "I want to accomplish X, Y, or Z: how can I do that?" This requires a new way of looking at biology – one that makes it

[18] OIE, Advances in Biotechnology and Future Impact on Animal Health, OIE Bulletin, Vol.4, 2007, http://www.oie.int/eng/publicat/BULLETIN%20PDF/Bull%202007-4-ENG.pdf

[19] Smith, Two studies highlight promise and problems for gene silencing technique, News, Nature, 26 March 2008.

[20] Geddes, New Gene Therapy Targets Cholesterol, New Scientist, 17 May 2007, http://www.newscientist.com/article/mg19426044.000-new-gene-therapy-targets-cholesterol.html

[21] Callaway, Hepatitis C is first target for new therapy, New Scientist, 26 March 2008, http://www.newscientist.com/channel/health/dn13539-hepatitis-c-is-first-target-for-new-therapy.html

[22] MicroRNAs Make Big Impression in Disease after Disease, Science, Vol. 319, 28 March 2008.

[23] GlaxoSmithKline does deal to develop MicroRNA Drugs, News in Brief, Nature, Vol. 452, 24 April 2008.

easier to handle. Drew Endy, one of the leading researchers in the field of synthetic biology, has detailed the necessary steps:

> "Synthetic biology builds on genetic engineering by adding three new foundational technologies: automated sequencing and synthesis (to make it easier to build things); standards (defining the things that are being built); and abstraction (to hide biological complexity). Each of these new technologies are just as important as any of the biotechnology techniques and taken together they define a new process for engineering biological systems. This is called synthetic biology"[24].

23. Some of these techniques are more developed (for example, see gene synthesis above for details on gene sequencing and synthesis). Others are well established principles that have yet to be acted upon (for example, efforts to mask the complexity of biology are not new, yet it is still taught the same way)[25]. The remainder are only in their infancy (for example, standardisation – such as Endy's BioBricks Foundation for the standardisation of biological parts[26]).

24. It is clear that all of the conditions for successful biological engineering are yet to be realised. However, it is possible to envisage them coming to pass in the medium to long-term. Given the potential impact of synthetic biology on the life sciences and everyday life, it also has a great potential for use for malign ends. Some of those involved in synthetic biology are already considering this issue, and how to manage potential risks.

IV. Open-source publication of raw research data

25. Recent developments in information technology have significantly changed the way information is exchanged. A great deal of information available online is no longer centralised or static, but is dynamic and user-contributed. (For example, think of the difference of a traditional website and Wikipedia – it would be practically impossible for a single institution to maintain the online encyclopaedia).

26. According to a review in Scientific American, certain research centres around the world are opening up their research data, laboratory notes and previously unpublished information for online collaboration[27]. Advocates of publishing all raw data, thoughts, insights and suggestions online suggest that "when you do your work online, out in the open, you quickly find that you're not competing with other scientists anymore but cooperating with them". This increase in efficiency, they claim "could greatly benefit society, in everything from faster drug development to greater national competitiveness".

27. As the review notes, many scientists remain wary of such openness, especially in the fields most relevant to the Convention – "putting your serious work out on blogs and social networks

---

[24] Endy, Defining Synthetic Biology, YouTube, [Online] http://www.youtube.com/watch?v=XIuh7KDRzLk

[25] For the importance of abstraction, see: Endy, Adventures in Synthetic Biology, Nature, http://www.nature.com/nature/comics/syntheticbiologycomic/index.html

[26] See: http://bbf.openwetware.org/

[27] Waldrop, Science 2.0 – Is Open Access Science the Future? Scientific American, 21 April 2008.

feels like an open invitation to have your lab notebooks vandalised – or worse, your best ideas stolen and published by a rival".

28. The OpenWetWare project of the Massachusetts Institute of Technology is an example of this open approach[28]. The project uses the same technology used by Wikipedia and so, is maintained by its member community. It was developed originally as a better way to keep the relevant websites up-to-date. As more and more students became contributors and users, it became "a convenient place to post… lab techniques,… lore in biology labs [that] never makes it into protocol manuals,… [and] information useful to the lab's members".

29. Access to open-source research data, albeit in a more traditional setting, has already yielded some practical results. For example, one research team identified and published one of the main biochemical pathways for drug addiction without doing a single experiment[29]. Instead, they mined data out of existing and published research data. The more information publicly available, the more that can be done with it.

30. This too, however, could have implications for the Convention. Being able to extract new meaning by combining old data in new ways has the potential for malign as well as beneficial application. The more raw data available, the greater the risk it could be used in contravention to the aims of the Convention. Furthermore, raw data might possess a greater chance of misuse than traditional published data. Published data will almost exclusively relate to a beneficial purpose, for example ways to increase the efficacy of an antibiotic. Raw data could well include information that produces precisely the opposite effect than desired (in this case reducing the efficacy of the antibiotic). While it is unlikely this data would make it out into the public domain through a traditional paper (it does not support the main hypothesis and contradicts the stated aim of the study), it would be released if all data derived from the experiment were to be made available online.

———

[28] For more information see: http://www.openwetware.org

[29] Drug Addiction: Going by the Book, The Economist, 10 January 2008, http://www.economist.com/science/displaystory.cfm?story_id=10493159