

Cryptography techniques to screen synthetic DNA could help prevent the creation of dangerous pathogens

October 5 2018, by Rob Matheson



Kevin Esvelt directs the Sculpting Evolution Group at MIT Media Lab. Credit: Maciek Jasik

In 2016, synthetic biologists reconstructed a possibly extinct disease,

known as horsepox, using mail-order DNA for around \$100,000. The experiment was strictly for research purposes, and the disease itself is harmless to humans. But the published results, including the methodology, raised concerns that a nefarious agent, given appropriate resources, could engineer a pandemic. In an op-ed published today in *PLOS Pathogens*, Media Lab Professor Kevin Esvelt, who develops and studies gene-editing techniques, argues for tighter biosecurity and greater research transparency to keep such "information hazards"—published information that could be used to cause harm—in check. Esvelt spoke with MIT News about his ideas.

Q: What are information hazards, and why are they an important topic in synthetic biology?

A: Our society is not at ease with this notion that some [information](#) is hazardous, but it unfortunately happens to be true. No one believes the blueprints for nuclear weapons should be public, but we do collectively believe that the genome sequences for viruses should be public. This was not a problem until DNA synthesis got really good. The current system for regulating dangerous biological agents is bypassed by DNA synthesis. DNA synthesis is becoming accessible to a wide variety of people, and the instructions for doing nasty things are freely available online.

In the horsepox study, for instance, the information hazard is partly in the paper and the methods they described. But it's also in the media covering it and highlighting that something bad can be done. And this is worsened by the people who are alarmed, because we talk to journalists about the potential harm, and that just feeds into it. As critics of these things, we are spreading information hazards too.

Part of the solution is just acknowledging that openness of information has costs, and taking steps to minimize those. That means raising

awareness that information hazards exist, and being a little more cautious about talking about, and especially citing, dangerous work. Information hazards are a "tragedy of the commons" problem. Everyone thinks that, if it's already out there, one more citation isn't going to hurt. But everyone thinks that way. It just keeps on building until it's on Wikipedia.

Q: You say one issue with synthetic biology is screening DNA for potentially harmful sequences. How can cryptography help promote a market of "clean" DNA?

A: We really need to do something about the ease of DNA synthesis and the accessibility of potential pandemic pathogens. The obvious solution is to get some kind of screening implemented for all DNA synthesis. The International Gene Synthesis Consortium (IGSC) was set up by industry leaders in DNA synthesis post-anthrax attacks. To be a member, a company needs to demonstrate it screens its orders, but member companies only cover 80 percent of the commercial market and none of the synthesis facilities within large firms. And there is no external way to verify that IGSC companies are actually doing the screening, or that they screen for the right things.

We need a more centralized system, where all DNA synthesis in the world is autonomously checked and would only be approved for synthesis if harmful sequences were not found in any of them. This is a cryptography problem.

On one hand, you have trade secrets, because firms making DNA don't want others to know what they're making. On the other hand, you have database of hazards that must be useless if stolen. You want to encrypt orders, send them to a centralized database, and then learn if it's safe or

not. Then you need a system for letting people add things to the database, which can be done privately. This is totally achievable with modern cryptography. You can use what's known as hashes [which converts inputs of letters and numbers into an encrypted output of a fixed sequence] or do it using a newer method of fully homomorphic encryption, which lets you do calculations on encrypted data without ever decrypting it.

We're just beginning to work on this challenge now. A point of this PLOS Pathogens op-ed is to lay the groundwork for this system.

In the long term, authorized experts can add hazards to their own databases. That's the ideal way to deal with information hazards. If I think of a sequence that I'm confident is very dangerous, and people shouldn't do this; ideally I would be able to contribute that to a database, possibly in conjunction with just one other authorized user who concurs. That could make sure nobody else makes that exact sequence, without unduly spreading the hazardous information of its identity and potential nature.

Q: You argue for peer review during earlier research stages. How would that help prevent information hazards?

A: The horsepox study was controversial with regard to whether the benefits outweighed the risks. It's been said that one benefit was highlighting that viruses can be built from scratch. In oncological viral therapy, where you make viruses to kill cancer, [this information] could accelerate their research. It's also been postulated that horsepox might be used to make a better vaccine, but that the researchers couldn't access a sample. Those may be true. It's still a clear information hazard. Could that aspect have been avoided?

Ideally, the horsepox study would have been reviewed by other experts, including some who were concerned by its implications and could have pointed out, for example, that you could have made a virus without harmful relatives as an example—or made horsepox, used it for vaccine development, and then just not specified that you made it from scratch. Then, you would have had all the research benefits of the study, without creating the information hazard. That would have been possible insofar as other experts had been given a chance to look at the research design before experiments were done.

With the current process, it's typically only peer review at the end of the research. There's no feedback at the research design phase at all. The time when peer review would be most useful would be at that phase. This transition requires funders, journals, and governments getting together to change [the process] in small subfields. In fields clearly without information hazards, you might publicly preregister your research plans and invite feedback. In fields like synthetic mammalian virology that present clear hazards, you'd want the research plans sent to a couple of peer reviewers in the field for evaluation, for safety and for suggested improvements. A lot of the time there's a better way to do the experiment than you initially imagined, and if they can point that out at the beginning, then great. I think that both models will result in faster science, which we want too.

Universities could start by setting up a special process for early-stage [peer review](#), internally, of gene drive [a genetic engineering technology] and mammalian virology experiments. As a scientist who works in both those fields, I would be happy to participate. The question is: How can we do [[synthetic biology](#)] in a way that continues or even accelerates beneficial discoveries while avoiding those with potentially catastrophic consequences?

More information: Kevin M. Esvelt et al. Inoculating science against

potential pandemics and information hazards, *PLOS Pathogens* (2018).
[DOI: 10.1371/journal.ppat.1007286](https://doi.org/10.1371/journal.ppat.1007286)

This story is republished courtesy of MIT News (web.mit.edu/newsoffice/), a popular site that covers news about MIT research, innovation and teaching.

Provided by Massachusetts Institute of Technology

Citation: Cryptography techniques to screen synthetic DNA could help prevent the creation of dangerous pathogens (2018, October 5) retrieved 21 June 2024 from <https://phys.org/news/2018-10-cryptography-techniques-screen-synthetic-dna.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.