US report pins down future biosecurity

Committee recommends a sequence-based system for identifying pathogens.

Can the disease-causing capabilities of an organism be predicted from its DNA? This was a key question faced by a 13-member committee of the US National Research Council (NRC). It was trying to determine what it would take to develop a government system that spots bioweapons in the making by screening the genetic sequences routinely ordered from commercial suppliers of synthetic DNA.

This week, the committee offered its answer in a 187-page report commissioned by the National Institutes of Health (NIH). The verdict: a biosecurity system that can predict the potential for harm lurking within a snippet of DNA is so technologically distant that the concept is useless for practical purposes. "This is a prediction problem that can't be solved, now or in the foreseeable future," says Sean Eddy, one of the report's authors and a computational biologist at the Howard Hughes Medical Institute's Janelia Farm Research Campus in Ashburn, Virginia.

The committee also declined to describe a detailed scientific road map leading to such a predictive capability, because it felt that the information could be misused. "We were very hesitant to go down that path, because we felt that the ability to predict something relied on the same skill set that would be needed to design a pathogen," says James Leduc, chair of the committee and director of the Galveston National Laboratory at the University of Texas Medical Branch in Galveston.

The NRC report comes less than three months after Craig Venter and his colleagues at the J. Craig Venter Institute in Rockville, Maryland, published their manufacture and insertion of a synthetic bacterial genome into a closely related bacterial cell which was then able to self-replicate (D. G. Gibson *et al.* *Science* **329**, 52–56; 2010). This milestone lifted the profile of synthetic biology, including its potential for misuse.

Prompted by advances such as this, the committee did, however, identify a key change that would be possible with current technology: moving to a sequence-based classification system for the regulation of dangerous pathogens. The United States regulates a list of 82 pathogens and toxins, called 'select agents', deemed to pose a biosecurity threat and so subject to restricted access. But currently, nothing identifies them beyond taxonomic labels, such as *Bacillus anthracis* for anthrax.

D. SCHWARTZ/ISTOCKPHOTO

All in the code: is it possible to predict the pathogenicity of a toxin by looking for particular genes?

"A sequence-based classification system," says the report, "could be used to create a pragmatic 'brighter line' for deciding when a new genome sequence should be regarded as one of the existing select agents or not." This would help, it suggests, to tackle potential confusion raised by variants of existing pathogens. The report states that DNA synthesis companies, and the scientists they serve, should be able to quickly and unambiguously determine whether a given sequence is on the select-agent list.

The report also describes a "yellow flag" biosafety system that would address sequences of concern — snippets of DNA that are not in themselves select agents, but could be part of one or otherwise used to produce a bioweapon. The yellow-flag system would consist of a centralized biosafety sequence database that would be annotated as evidence of the function of suspect genes comes to light. If a sequence received a yellow-flag designation, that should not trigger regulatory action, the authors write, but "common sense follow-up", such as a telephone call from a synthetic DNA company to make sure that a customer is legitimate.

The idea of relying on sequences to define select agents drew some praise. "It's very good because it places a lot of emphasis on really using sequences to screen, rather than worrying about taxonomy," says Jeremy Minshall, president of DNA 2.0, a synthetic-genomics company in Menlo Park, California.

Currently, he says, to comply with select-agent

regulations, companies like his must laboriously comb GenBank — an annotated database of publicly available DNA sequences maintained by the NIH's National Center for Biotechnology Information — for sequences that could correspond to a select agent on the list maintained by the Centers for Disease Control and Prevention in Atlanta, Georgia, and the US Department of Agriculture. "It's not a trivial task," says Minshall, who adds that he would welcome the efficiency of a comprehensive government curated database of pathogenic sequences.

But some critics say that moving to sequence-based classification would introduce complexity. "It would actually decrease regulatory clarity," says Gigi Kwik Gronvall, a senior associate at the Center for Biosecurity of UPMC, a major hospital network in Pittsburgh, Pennsylvania. "It exchanges a functional definition of a pathogen — that is, a microorganism that can do harm — for a very complicated approach that says: 'Maybe it's a pathogen and maybe it's not, and there are infinite numbers of possibilities.'"

Eddy calls Gronvall's criticism "totally fair". But, he adds, it does not answer the question: "What do we do with the select agent list in an era when you can synthesize things for which there are no experimental data?"

The NIH did not have immediate comment on the report, which stops short of saying that the government should press ahead with building the suggested classification system. Committee members wrote that they chose to leave that decision — including the risk-benefit analysis that should precede it — to policy-makers. ■

Meredith Wadman

"It places a lot of emphasis on really using sequences to screen, rather than worrying about taxonomy."