EBRC Engineering Biology Research Consortium

Platform vulnerabilities and security in the bioeconomy

A Policy Paper by the Engineering Biology Research Consortium

Compiled and edited by Becky Mackelprang, EBRC Associate Director for Security Programs

December 2022

The foundational tools and technologies of engineering biology developed over the past 20 years are enabling the development of commercial products for a rapidly expanding bioeconomy. Many developing applications of engineering biology rely on platforms such as gene editing. Platform technologies "typically are highly shared, have multiple purposes and uses, and offer tremendous benefits through their scalability and adaptability."[1] However, because they can underpin products and systems across applications, they can "introduce new and shared vulnerabilities that can be exploited to misuse any technology on the platform." It is thus important to recognize the platform technologies and systems that underpin the bioeconomy, understand their vulnerabilities, and identify steps the US Government can take to prevent or mitigate their exploitation. Crucially, in the identification of such steps, USG must also recognize that security measures that prohibit areas of research, sharing of information, and collaboration across national boundaries can undercut the bioeconomy and cause real loss to the development of life- and planet-saving biotechnologies. Thus, decisions to implement security measures should account both for security benefits and lost opportunities.

Below, a non-comprehensive view of platform technologies, their vulnerabilities, and potential exploits of those vulnerabilities is described. Platforms are divided into three sections—biology-based platforms, automation platforms, and computation and data-based platforms—with an additional section to consider vulnerabilities and exploits that are relevant across platforms. Much of this information is also available in the linked table, which includes segments of the bioeconomy that particularly rely on given platforms.

Platform technologies, vulnerabilities, and exploits specific to the bioeconomy

All platforms

- An overarching vulnerability across platforms is their very existence and the potential for another nation to develop new or improved platforms that are widely adopted. If another nation is able to offer superior, cost-competitive DNA synthesis or sequencing capabilities, lab automation systems, or computational tools, that nation could achieve platform dominance. This would come at an economic loss to the U.S. That economic loss could be amplified many fold if the competitor nation uses the data from platform users to bolster their own research and innovation efforts.
- Insufficient regulatory capacity for a complex regulatory landscape: If regulatory agencies do not have the necessary staffing and resources to handle the coming tidal wave of products developed on biotechnology platforms, the U.S. may miss seeing the return on its investments in research and development. If a product lacked sufficient review and caused damage to a consumer because of an insufficient or overburdened regulatory process, entire segments of the bioeconomy could be shut down. And, when regulations are too complex and/or stringent, we forfeit US competitive advantage and leadership across the global bioeconomy.

[1] National Academies of Sciences, Engineering, and Medicine 2022. Protecting U.S. Technological Advantage. Washington, DC: The National Academies Press. https://doi.org/10.17226/26647.

Analytical and biology-based platforms

The research and development required for innovative biotechnologies is enabled by biologically-based platforms and the analytical platforms that enable measurement and analysis of biological activity and outputs. For example, gene sequencing, gene editing, and DNA synthesis are biological-based platforms while proteomic and metabolomic platforms enable the measurement of biological systems. These platform technologies enable the development of commercializable products across segments of the bioeconomy from medicine to agriculture to consumer cosmetics and fashion. Many of these platforms share common vulnerabilities, including:

- Disinformation: With many consumers still carrying a negative view of the products of bio-innovation, such as genetically engineered foods, bio-based platforms such as gene editing are vulnerable to disinformation. This could be exploited by adversarial groups or nations seeding and/or amplifying disinformation that destroys trust in products of the bioeconomy.
- Introduction of contaminating or counterfeit chemicals or infectious agents to source materials for the bioeconomy: Biology-based platforms depend on high-quality, consistent reagents and feedstocks. The interception and tampering of a shipment e.g. of nucleotides en route to a DNA synthesis company—or worse yet tampering by an insider post quality control measures—could impact tens of thousands of orders and researchers, perhaps resulting in a lack of primers for medical diagnostic PCR tests, halting research, and delaying the biological parts and services supply chain. Such exploits should be viewed with acknowledgement that accidental contamination and supply chain issues caused by natural disasters, pandemics, and global political realities are more likely and may have broader impacts.

Automation and scale-up platforms

Engineering biology researchers often employ a design - build - test - learn (DBTL) cycle to develop systems and organisms that are optimized for a given function or outcome. It is not uncommon for institutions to invest in automation platforms, which enable more samples or experiments (and thus DBTL cycles) to be run in a given time, enhance reproducibility, and save on labor costs. And, as biological systems are optimized, they must be scaled up to commercial levels. Scaling and manufacturing facilities and platforms are key nodes whose vulnerabilities may be ideal targets for nefarious actors.

- The code that runs automation platforms or pieces of automation equipment may be vulnerable to hacking. Vendors may or may not commit the resources required to penetration-test their own code. Furthermore, users may buy pieces of equipment from different vendors and develop their own code to enable communication between instruments. Such users generally focus their code development on functionality without dedicating sufficient resources to security. Thus, vulnerabilities in the code driving a single piece of equipment or connecting several pieces of equipment could be exploited by an outsider (or insider) gaining access and, for example, undetectably changing the automated process to give inaccurate results.
- Scale-up platforms are susceptible to contamination and physical infrastructure damage. Microorganisms could be introduced to massive bioreactors, inhibiting the production of or contaminating desired products. Physical infrastructure damage could cause leaks of engineered organisms or toxic by-products.

Computation and data-based platforms

Biotechnology innovation is enabled and informed by advanced computational capabilities and platforms. Huge searchable databases such as GenBank®, which contained 240,539,282 sequences and 1,562,963,366,851 bases as of Oct 2022,[2] provide up to date, comprehensive DNA sequence information across the scientific community. Computing platforms enable researchers across application areas to process "big data" rapidly, hastening discovery and shortening the DBTL cycle. Machine learning and large scale modeling are improving significantly each year and are close to reliably predicting protein structure and function from sequence. Furthermore, researchers are working on an even bigger challenge — generating sequences based on the input of a desired function, which could drastically reduce research timelines. Vulnerabilities of these computation and data-based platforms include:

- Imperfect taxonomic labeling in searchable databases: When researchers deposit a synthetic sequence in a database—for example, a single sequence containing DNA from both Ebola and benign Green Fluorescent Protein—that sequence has to ultimately be given a single taxonomic label. That taxonomic label has implications for other researchers and for security processes that reference these databases, such as DNA synthesis screening. One can imagine purposely feeding such databases inaccurate information which could enable concerning sequences to escape screening by DNA synthesis providers or, more broadly, result in experiments premised on inaccurate data. This would require significant knowledge of how these databases employ quality control and would not be a trivial pursuit.
- Lack of adequate code security:
  - Advances across life sciences have made it more common for researchers to collect gigabytes, or even terabytes, of experimental data. Because not all researchers are competent in handling and processing such large swaths of data, individual researchers and companies may use computational analytic platforms. Developers of such platforms—whether commercial or academic—may not have the funding, awareness, and/or motivation to implement appropriate cybersecurity measures, potentially leaving the platform and its users susceptible to attacks that compromise system fidelity and result in inaccurate experimental results and analysis.
  - The ability to generate sequence information for a desired protein function (e.g., conversion of a common metabolite to a therapeutic metabolite) would be a powerful platform technology. It could be exploited to develop methods for biologically producing chemical compounds with high human toxicity. This exploit would be very challenging, requiring significant time, expertise, infrastructure, and trial and error. Compounds with high human toxicity also often kill the microbial cells one may try to produce them in (see Urbina et al, 2022[3]).

Addressing vulnerabilities

Some of these vulnerabilities arise from the same characteristics that enable innovation in the bioeconomy, such as the sharing of tools and information. Thus, those making efforts to address vulnerabilities should be cognizant that regulatory interventions may present significant risk of ceding global leadership of the

[2] NIH National Library of Medicine, National Center for Biotechnology Information. GenBank and WGS Statistics (Accessed Nov 17, 2022). https://www.ncbi.nlm.nih.gov/genbank/statistics/

[3] Urbina, F., Lentzos, F., Invernizzi, C. et al. Dual use of artificial-intelligence-powered drug discovery. Nat Mach Intell 4, 189–191 (2022). https://doi.org/10.1038/s42256-022-00465-9.

bioeconomy, which is a vulnerability itself. The United States Government can minimize platform vulnerabilities across the bioeconomy through steps such as:

- Conducting outreach to large and growing bioeconomy platform and equipment vendors to educate companies on the importance of cyber- and physical security practices, insider threat programs, and process-monitoring to detect intrusion and sabotage attempts.
- Supporting public/private partnerships that work with platform companies to detect vulnerabilities through tabletop exercises, red-teaming, etc., and are trusted to rapidly disseminate warning messages and patch/mitigation strategies to user communities.
- Encouraging platform companies to develop and implement continuity of operations plans and capabilities to ensure robustness in the face of significant temporary supply chain disruption, natural disaster, etc., and to be capable of delivering when surge production is necessary.
- Dedicating resources to the characterization of these vulnerabilities so that i) those that carry the greatest risk can receive the greatest attention and ii) mitigation strategies can be developed with stakeholder input that do not compromise U.S. leadership of the global bioeconomy.
- Supporting workforce development so that the workforce operating these platforms are well-trained, which will make them better able to understand and implement safety and security measures.
- Supporting the incorporation of safety and security into education and workforce training so that participants in the bioeconomy understand why safety and security are important and how to implement best practices.
- Developing and perpetuating standards for interoperability, accountability, measurability, safety, and security in engineering biology and the bioeconomy. For example, the United States is the only nation to provide screening guidance to providers of synthetic DNA. Some international companies adhere to screening standards in compliance with their nation's export laws or as required for membership in the International Gene Synthesis Consortium, but efforts that encourage and support the global adoption of screening practices would support the continued productivity of that platform for peaceful research and innovation. Similarly, the US might consider establishing leadership in efforts such as the development of metrics and standards around the bioeconomy or a unified standard for obtaining, storing, and disseminating genetic information.

Many platform technologies underpin the bioeconomy and enable incredible discovery and innovation across its segments. However, these platforms can also present vulnerabilities, and the USG must place these vulnerabilities in context and estimate the risk of their exploitation. The measures suggested above, meant to secure and safeguard the bioeconomy and biotechnology without compromising innovation, will help us ensure that platform technologies in engineering biology can continue to help us address our nation's and planet's most significant challenges.