Blue Ribbon Panel Remarks

Hello my name is Corey Hudson. I am a manager of Computational Biology and Biophysics at Sandia National Labs. Chris was not able to make this meeting today. As I will discuss later in this talk, Chris and I are part of a multi-laboratory team working on biological cybersecurity and cyberbiosecurity, involving Chris lab - PNLL, my lab - Sandia, Los Alamos and several other collaborators in government, industry and academia. The remarks that follow are Chris' but I am happy to field any questions for me or the panel during the discussion that follows.

I would like to begin by thanking the blue ribbon study panel, including Governor Tom Ridge, Senator Tom Daschle, Representative Jim Greenwood, and The Honorable Ken Wainstein for the opportunity to participate to this panel today. I truly believe that the work being done by the Blue Ribbon Study Panel is of utmost importance in addressing an urgent and compelling concern facing the U.S.—namely the increasing risk to our bioeconomy through cyber and physical threats.

Let me begin by defining cyber-bio security through the lens of a scientist who bridges both bioinformatics and cyber as this definition is broader than some often implied by this term. Our team represents a multi-laboratory team that also includes Pacific Northwest National Laboratories, Sandia National Laboratories, and Los Alamos National Laboratory as well as Randy Murch at Virginia Tech and Sp. Agent Ed You, who will be addressing this panel later today. Together we have defined cyber-bio security as a broad collection of activities with stakeholder communities focused on collectively ensuring the security and integrity of national resources in biotechnology. This includes the familiar areas of national biothreat countermeasure stockpiles and biological disaster response infrastructure. But it goes far beyond this. Much of our biotechnology industry, academic and government research depends on a global, communal ecosystem of data, software, and computing infrastructure—all of which resides in a hostile cyber environment accessible to adversaries in all regions of the globe. This public resource began as a well-intentioned global science community forum to facilitate rapid advancement through sharing of data and software, enabled by a growing availability of computing on demand through technologies like cloud architectures and an aggressive effort to make federally funded research openly available. Security was not a primary concern because scientific credibility compelled users to exert great efforts to share only high quality data and applications. But as this resource grew in size and usefulness, so too did its impact on public health, policy, and economic impact on our biotechnology and big pharma industrial sectors—as well as DoD and national readiness posture. Today there are many gigabytes of publicly accessible and publicly changeable data that we use every day to develop new biotechnologies. There are also computer-driven genome sequencing systems and downstream analysis platforms such as those maintained by the DOE Joint Genome Institute that together are producing so much new data that the total body of biological data stored on computing resources doubles approximately every 18 months outpacing growth in analytical computing resources themselves. Vaccines and medicines are being designed using this and other data, along with a host of publicly generated and open access software tools, sometimes running on public cloud platforms. All of this has driven a modern revolution in the rate of new biological insight which relies on the public, interactive nature of this critical infrastructure. And the intersection of this with the Internet of Things revolution will increase efficiency of the whole system, dramatically expanding its digital footprint—all with unknown security implications.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Now let us imagine what is at risk if this resource were compromised using the classical cyber security triad of Integrity, Availability, and Confidentiality. From an Integrity perspective, if data in sequence repositories or sequence analysis centers is corrupted—regardless of whether this is malicious or accidental— then vaccines could be created that attack the wrong virus, or no virus at all, leaving us unprotected during a biological disaster. Technologies that identify biological agents could give incorrect answers, blinding us to the presence of threats and delaying response at the cost of lives. Further, non-viable candidate medicines could be developed at great cost to our biotech industry while other nations pass us by, developing viable medicines and capturing our competitive advantage. From an Availability perspective, if the nation were denied access to these resources at a critical moment, for instance while we are attempting to respond to a never-before-seen biological threat, we might lose our ability to detect and respond, again at the cost of lives. And from a Confidentiality perspective, losing privacy of our biological data risks exposing the personal information of the brave people such as the participants in the Veteran's Administration 1 Million genome project, who have so graciously provided their biological data so that we may all benefit.

Today, so much of our safety and economic well-being relies on the security of this resource, that it is "too big to fail". Yet without security being designed in from the beginning, we are left vulnerable to deliberate or accidental misuse that can have nationally catastrophic effects. Given this understanding, it is our recommendation that the nation:

- First, clearly define roles, authorities, and expectations for departments and agencies that own various pieces of the solution to this challenge;
- Second, prioritize and fund research into novel solutions to safeguard bioeconomy-protecting resources without limiting the utility afforded in the current environment; and
- Third, create bridges between the many stakeholder groups including academia, federally funded research and development centers, industry and vendors, and government institutions to facilitate exchange of ideas, technology, and practical solutions.

It is essential that whatever we do to protect this resource does not interfere with its usability because any slowdown in the process risks our position at the global forefront of innovation in biotechnology as well as our biodefense readiness.

Defining clear roles, authorities and expectations for government is a complex challenge not unlike the challenges this Panel has already addressed. Many agencies with different bioscience and biodefense equities are involved, including:

- DOD, which has responsibilities for biodefense, as well as warfighter readiness and veterans' health;
- HHS, which is responsible for human health and most of the associated basic science research, and maintains many of the pertinent databases;
- DHS, which is responsible for the nation's critical information technology infrastructure;
- DOE, which shepherds the Joint Genome Institute and its informatics tools, and has expertise in development of data analytics and massive and secure data processing and management; and
- the Department of Justice and the FBI, which are responsible for the collection, analysis, and dissemination of cyber threat information and related law enforcement activities.

In addition to the US entities, there are significant foreign resources that provide significant value to the bioeconomy, such as UniProt, the Universal Protein Resource, which is a European entity. These entities should also be encouraged to work to ensure the security and integrity of their resources.

It is imperative that the nation also invests in research into novel solutions to safeguard the bioeconomy. There are many emerging technology areas that may have a central role in securing the bioeconomy going forward. Machine learning and artificial intelligence make it possible to improve on the efficiency of human decision making and can speed up discovery. But transparency into these approaches is needed as well as better understanding of how they can be manipulated. Large-scale computing, including computing at the edge which will likely be accelerated by the coming 5G telecom revolution, drives development of smart devices, data collection platforms, and can even steer scientific instruments in real time. But a way must be found that allows this to happen securely on untrusted systems being used by users with unknown intent. Data sciences including information fusion and statistical and other mathematical techniques applied at all stages of the data/discovery cycle can help quantify our uncertainty in results which is essential for good decision making. But how do we ensure the accuracy and integrity of this data in a hostile computing environment? There are many other technologies and R&D areas that could potentially result in substantial improvements to the security of the biotech research environment.

Finally, as the Panel knows well, stakeholder community engagement will be essential to the development of robust cyber-bio security solutions. Academia, FFRDCs, government, and industry partners all have different interests and investments in our publicly available resources. Communication among these stakeholders can help accelerate the understanding of the features of the real issues, and continued engagement throughout the development of solutions will be required to increase the likelihood that end users will actually adopt new security technologies. If these stakeholders can be brought together to determine requirements for practical solutions, the research and development community and ultimately, the industrial community, can plan and execute roadmaps to realize these solutions at a reasonable cost.

In closing, I would like to summarize by reiterating that protecting our bioeconomy is an urgent matter with far-reaching impacts in public health and well-being, readiness posture for responding to biothreats, and maintaining the competitiveness of the US in the global economy. The ecosystem of biotechnology that underpins this resource is, by design, accessible and highly dynamic to the pace of modern biology research. Fueled by continuous improvements in biotechnology, availability of computing, data networking, open source software development, and smart networked devices, our bioeconomy is at the forefront of importance for ensuring the American way of life. Yet if we don't prioritize discovering ways to protect this resource without limiting its utility, we are vulnerable. The good news is that many institutions are poised to do their part in collectively solving this problem. The better news is that the Blue Ribbon Study Panel is bringing this important issue to the forefront of the national dialogue. As a scientist, I applaud the Panel for tackling this technically challenging but critical issue. With your help in developing a clearly illuminated path forward, I am confident that we will rapidly advance our ability to defend and protect the US bioeconomy, thereby enabling a more secure future.