INNOVATING CROSS-DOMAIN SOLUTIONS TO DETECT EMERGING BIOLOGICAL THREATS

Sharing Accurate And Secure Pathogen Genome Data

Sterling Sawaya GeneInfoSec Inc.

The ability to detect or predict biological threats involves the use of genetic data. Often this genetic data is communicated from international partners. These partners can have agendas that do not necessarily align with the US, and may rightfully fear that the data they share can be misused. Potential misuse of pathogen data can be as simple as a failure to reciprocate data sharing, or a circumstance in which value obtained from the data is not shared (e.g. from profitable vaccines). Other serious misuse of the data can also occur, in which insecure pathogen data is used to design novel bioweapons. An immediate concern is that pathogens can be accurately replicated with synthetic DNA, so the sharing of full pathogen genome data is, in effect, providing the ability to reproduce the organism. This situation would lead to an inability to attribute any future malicious use the organism. This challenge is especially difficult for partners that expect to encounter potential bioweapons, such as countries with former Soviet bio-labs. Even if these countries recognize a need to share the data, their own internal policies may forbid it. We have developed a solution to this problem, in which data is shared in a format that does not allow full genomes to be reconstructed. By scrambling multiple genomes together, adding decoy sequence data, and potential masking some regions of the genome, raw pathogen sequence information can be shared while entirely avoiding the potential for full genome reconstruction. Such data can still be used to monitor pathogens, determine their risk, and if needed develop countermeasures. Through software, our approach can be applied to any existing genetic data today. When considering data security, however, we also must consider the source of the data, the data's integrity, and whether the data has remained confidential. To provide 100% assurance that genetic data is secure, we can provide an additional layer of security. Through molecular cryptography we can protect data within molecules in a test-tube, securing genetic data before it ever touches a potentially insecure network. Within DNA molecules we can apply the same techniques as in our software. Data from multiple sources can be scrambled together, decoy data can be added, and regions can be masked. The resulting pool of encrypted molecules can then be securely sequenced on insecure networks, or the genetic material itself can be securely transported through hostile territory. Data can then be unlocked, in part or in whole. This talk will discuss the need for such technology, details about its current readiness level, and highlight specific potential applications. As we look to build multi-sectoral, multi-lateral partnerships for sharing pathogen information, we can ensure strong leadership by offering exceptional security. This new layer of security will allow our partners to control their sensitive, valuable pathogen data, while giving them the confidence to share the data.