INNOVATING CROSS-DOMAIN SOLUTIONS TO DETECT EMERGING BIOLOGICAL THREATS

Towards Sequencing For Agnostic Discriminative Detection Of Biological Threats With Outbreak Potential

Frederik Schulz Lawrence Berkeley National Laboratory **Tomas Tyml** Joint Genome Institute / Lawrence Berkeley National Laboratory **Eva Myskova** Joint Genome Institute, Lawrence Berkeley National Laboratory **Lorenzo Aureli** Joint Genome Institute / Lawrence Berkeley National Laboratory **Anastasiya Sakkos** Joint Genome Institute / Lawrence Berkeley National Lanboratory **Yuko Yoshinaga** Joint Genome Institute / Lawrence Berkely National Laboratory **Harshini Mukundan** Office of National and Homeland Security / Lawrence Berkeley National Laboratory **Nigel Mouncey** Joint Genome Institute / Lawrence Berkeley National Laboratory

The detection and characterization of pathogens in complex samples remains a significant challenge due to the limitations of established methods. Traditional approaches typically rely on nucleic acid amplification with highly specific primers, antibody-based assays or pathogen-specific cultivation techniques. Whilst these methods are highly specific they typically require prior knowledge of the target pathogen. Additionally, these methods are primarily based on sole detection of a pathogen and are only in some cases complemented with amplicon sequencing, which however, often yields only a small fraction of the pathogen genome. Resulting sequencing data has limited strain resolution and makes tracing emerging variants difficult, especially after genomic recombination events. Furthermore, it hampers functional risk assessment, which is crucial for understanding the potential impact of detected pathogens.

To overcome these challenges, we are developing a reproducible framework for unbiased, pathogen-agnostic detection and characterization of biological threats with outbreak potential directly from complex environmental samples, such as wastewater, but also clinical samples, such as human bodily fluids. In the first phase of the project, we are employing qPCR and shotgun metagenomic sequencing to identify the limits of detection for major groups of pathogens ranging from RNA and DNA viruses to bacteria and eukaryotic parasites. With the goal to enhance detection sensitivity for broad groups of pathogens, we combine, streamline and optimize existing enrichment approaches at both the particle and sequence level. An integral component of our workflow is high-throughput sequencing, both short and long-read based, as it enables sufficient genome recovery for functional prediction and variant analyses. To distinguish pathogens from the baseline microbiome, we employ an ensemble of microbial lifestyle prediction tools and in-house developed classifiers that accurately predict sequences linked to pathogenicity from metagenomic data.

Based on our initial benchmarking results, our assay will complement existing efforts for specific pathogen monitoring. It will be particularly powerful for early detection of emerging pathogens in difficult sample matrices and for tracking the temporal dynamics of pathogen populations in complex samples, providing valuable insights into the emergence, real-time evolution, and spread of pathogens. By eliminating the need for prior knowledge of target pathogens and enabling strain-level resolution, our pathogen-agnostic detection and characterization framework not only enhances national security through less biased pathogen detection but also allows for a better understanding of functions, including novel virulence factors, that are associated with novel pathogens. The insights gained from applying our assay to critical samples will have far-reaching implications for public health, environmental monitoring, and the development of targeted interventions to mitigate the impact of novel pathogens.

Approved for public release; distribution is unlimited.