INNOVATING CROSS-DOMAIN SOLUTIONS TO DETECT EMERGING BIOLOGICAL THREATS

Cas9-assisted Enrichment Of Metagenomic Sequencing Samples

Jessica Resnick The Johns Hopkins University, Applied Physics Laboratory, Laurel, MD **Kathleen Verratti** The Johns Hopkins University, Applied Physics Laboratory, Laurel, MD **Bryan Necciai** Joint Program Executive Office for Chemical, Biological, Radiological and Nuclear Defense (JPEO-CBRND), Joint Project Lead for CBRND Enabling Biotechnologies (JPL CBRND EB), Frederick, MD **Shanmuga Sozhamannan** Joint Program Executive Office for Chemical, Biological, Radiological and Nuclear Defense (JPEO-CBRND), Joint Project Lead for CBRND Enabling Biotechnologies (JPL CBRND EB), Frederick, MD and Joint Research and Development, Inc., Stafford, VA **Claire Bell** The Johns Hopkins University, Applied Physics Laboratory, Laurel, MD

Metagenomics-enabled surveillance is a powerful tool for identifying pathogens present in an environment1. However, due to a lack of taxonomic resolution stemming from low abundance in the sample or low coverage when sequencing, it is often difficult to distinguish closely related pathogenic and non-pathogenic organisms, complicating decision making2. In an environmental sample, genomes from many different species may be present and enriching for low abundance genomes of interest is challenging. Here, as a proof of concept, we use a Cas9 combined with catalytically inactivated Cas9 (dead Cas9 aka dCas9), to target homologous sequences flanking variable regions (e.g., PAIs, O-antigen clusters) from a species of interest (Vibrio cholerae) in order to enrich for all strains or serogroups present, regardless of the sequence size of the variable regions. Guides were designed and optimized to target highly conserved rjg and gmhD genes that flank a variable area of the genome that may be used to differentiate between serogroups. Cas9 is first used to cut double stranded DNA at the outer ends of the conserved rjg and gmhD genes. Then, biotinylated dCas9 is used to bind the conserved regions remaining from each gene on the excised fragments. Streptavidin beads in turn bind to the biotin, which pulls down the whole complex of Cas9 including the unknown variable sequences, leaving behind unwanted sample DNA3. From the pulled down sequences, Illumina or Oxford Nanopore sequencing libraries may be prepared, significantly enriching the sequences of interest from the original metagenomic sample. The power of this approach is that it is not subject to the same bias towards fragment size as other pulldown methods and can be more easily multiplexed to multiple targets than traditional PCR methods. Additionally, obtaining strain (or serogroup) level information from environmental surveillance will allow for better epidemiological tracking, prediction of countermeasure efficacy, and attribution during outbreaks. The method described here can be further compared to Oxford Nanopore Technologies based adaptive sequencing and can also be applied to enrich all bacteria present in metagenome samples by targeting 16S rDNA regions.

References:

Ko, K.K.K., Chng, K.R. & Nagarajan, N. Metagenomics-enabled microbial surveillance. Nat Microbiol 7, 486–496 (2022). <a href="https://doi.org/10.1038/s41564-022-01089-w">https://doi.org/10.1038/s41564-022-01089-w

Shen, J. et. al. Toward Accurate and Robust Environmental Surveillance Using Metagenomics. Front. Genet. (2021) <a href="https://doi.org/10.3389/fgene.2021.600111">https://doi.org/10.3389/fgene.2021.600111

Gilpatrick, T., Lee, I., Graham, J.E. et al. Targeted nanopore sequencing with Cas9-guided adapter ligation. Nat Biotechnol 38, 433–438 (2020). https://doi.org/10.1038/s41587-020-0407-5