INNOVATING CROSS-DOMAIN SOLUTIONS TO DETECT EMERGING BIOLOGICAL THREATS

Development Of The Cascade Crispr Detection Guide Rna Design Pipeline

Colin Price MRIGlobal **Phil Davis** MRIGlobal **Alan Shteyman** MRIGlobal **Brian Clark** MRIGlobal **Joe Russell** MRIGlobal **Elaine Bradford** MRIGlobal **Julie Lucas** MRIGlobal **Richard Winegar** MRIGlobal

The use of CRISPR as a molecular detection system is emerging as a viable alternative to established methods such as PCR. CRISPR detection assays may be a more desirable technology in some applications relative to PCR, such as biosurveillance. CRISPR detection has higher specificity due to lack of non-specific amplification, higher sensitivity due to requiring less genetic material for detection, and quicker time-to-completion without requirement of time consuming amplification cycling. However, much like PCR primers, specific guide RNA sequences must be developed per genomic target. The Cas-CRISPR Automated Design and Evaluation (CasCADE) bioinformatics pipeline is a first of its kind guide RNA (gRNA) designer that allows for the selection and filtering of key criteria to include inclusion and exclusion groups, Cas protein type, scaffold sequence, PAM, and k-size. From a set of input genomic files, an inner set of kmers is first found that is conserved across all input records. If the design Cas protein type has a PAM site, conserved kmers are required to co-occur with the PAM region on the appropriate 5’ or 3’ end. The conserved candidates are then structurally evaluated with the addition of the scaffold sequence associated with the Cas protein type. Secondary structure protein folding algorithms are used to evaluate the stability of each candidate, and the free energy of each value is recorded as well. The best candidates are retained based on the structural stability, free energy, and sequence GC content. Each andidate is evaluated for inclusivity to the initial in group of submitted genomic files, as well as it’s exclusivity to either a user specified exclusive set of genomic files or the set of genomes from taxonomic near neighbors retrieved from NCBI. A final check is performed to determine if the assays are exclusive to humans using the GRCh38 reference genome. These factors are reported as a gRNA Quality Score which has been validated to accurately predict in vitro target detection success. CasCADE can design a CRISPR detection assay on the order of hours and scales elegantly to available hardware use, positioning CasCADE as a valuable asset in the biosurveillance space when quick turnaround times are paramount. To support the rise of CRISPR-based detection technologies, CasCADE provides an automated bioinformatics approach for the design and evaluation of detection gRNAs.

Support from the DARPA Biological Technologies Office as part of the Detect It with Gene Editing Technologies (DIGET) program funded under the Naval Information Warfare Center contract N66001-21-1-4048 which is awarded to MRIGlobal. The authors thank Craig Willis, Pamela Winegar, Landon Adebiyi, and Sarah Pope for programmatic support.