RAPID DETECTION OF ALL KNOWN MICROBES USING A MICROBIAL DETECTION ARRAY. C.

Jaing, S. Gardner, K. McLoughlin, J. Allen, J. Thissen, N. Be, T. Slezak, Lawrence Livermore National Laboratory (7000 East Ave., Livermore, CA, 94550, jaing2@llnl.gov).

Molecular detection of microbes: Recent advances in genomic-based technologies have revolutionized the field of microbial ecology and their influence on infectious diseases in human [1]. PCR and DNA sequencing approaches have been widely used for pathogen detection and characterization. PCR assays are limited, in that only a single or few organisms can be investigated per assay, with potentially high false-positive rates. While DNA sequencing can identify a larger scope of organisms, current DNA sequencing analysis methods are lengthy, costly and require significant computational time, and there is a lack of bioinformatic tools to rapidly identify and quantify abundances of species identified in a sample. In an effort to improve high-throughput analysis and detection, we have developed an innovative microarray platform called the Lawrence Livermore Microbial Detection Array (LLMDA) to probe for all known microbiological agents for which whole genomes, segments and plasmid sequences are available [2].

The LLMDA: The LLMDA technology can be applied to large numbers of environmental and clinical samples in a highly sensitive, specific, and cost-effective fashion. The recently updated LLMDA contains 180,000 probes designed to detect genomic DNA and cDNA from all currently sequenced microbial pathogens, a total of 10261 species, including 4219 viral, 5367 bacterial, 293 archaeal, 265 fungal, and 117 protozoan species that were sequenced through June, 2013. This microarray targets both conserved and unique genomic regions of sequenced microbial strains. The automated data analysis algorithm, Composite Likelihood Maximization (CLiMax), is integrated with a web interface that enables LLMDA data analysis within 30 minutes.

LLMDA in human health applications: LLMDA was recently established as a potential diagnostic platform for identification of viral pathogens in human clinical samples [3]. As little as 5 input genome copies ($\geq$1,000 copies/ml) of BK polyomavirus were clearly detected by LLMDA in urine samples after phi29 amplification [3]. Additionally, the LLMDA successfully identified other viral agents such as human papillomavirus, human herpesviruses, enteroviruses, and adenoviruses in a variety of human sample types including nasal swabs, urine, stool, serum, and cerebrospinal fluid [3]. In 2010, the LLMDA identified a porcine circovirus in the Rotarix vaccine [4,5]. This was the first study to report an adventitious contaminating virus from childhood vaccines, which demonstrated that LLMDA is a sensitive and powerful technology for vaccine safety monitoring.

LLMDA in environmental monitoring: The LLMDA technology has been evaluated in microbial detection in environmental air and soil samples. LLMDA detected as little as 100 genome copies of *B. anthracis* in spiked environmental air and soil samples, similar in sensitivity to next-generation metagenomic sequencing [6]. In a collaborative study with NASA JPL, the utility of LLMDA in detecting bacteria from environmental samples collected at the ISS was evaluated. The LLMDA successfully detected *P. acnes*, *S. aureas*, *S. epidermidis* and *S. cohnii*, confirming previous results by sequencing.

The Livermore Metagenomics Analysis Toolkit (LMAT): LMAT is a collection of software tools designed to identify the genes and organisms present in a shotgun metagenomic sample [7]. LMAT maintains the most complete collection of microbial genomes and genes publicly available in a rapidly searchable form. The database covers 12,632 species and stores 116 gigabases of searchable genomic data, which is roughly 3 times larger than any other published searchable database available for metagenomic search. LMAT ensures consistently fast runtimes and maintains high accuracy by using strain-specific k-mers (where k is typically 20 nucleotides), searching for common elelments only once. LMAT has been used to analyze many terabases of metagenomic data and provide novel insights on organisms present that were not previously considered.

Summary: We have developed a suite of genomic and bioinformatic tools including the Lawrence Livermore Microbial Detection Array and Metagenomics Analysis toolkit, scalable and easily adaptable to support space biology and astrobiology. The tools are highly sensitive and specific for genomic analysis and molecular characterization of human health and microbiomes, and to profile microbial contents in the environment.

References:

[1] Shafquat A et al. (2014) *Trends in Microbiol* 22: 261-266. [2] Gardner SN et al. (2010) *BMC Genomics* 11: 668. [3] Erlandsson L et al. (2011) *PLoS ONE* 6: e22631. [4] Jaing C et al. (2011) *PDA J Pharm Sci Technol* 65: 668-674. [5] Victoria JG et al. (2010) *J Virol* 84: 6033-6040. [6] Be NA et al. (2013) *PLoS ONE* 8: e73455. [7] Ames SK et al. (2013) *Bioinformatics*.