A QUANTITATIVE REDUCED REPRESENTATION SEQUENCING (QRRS) OF GENOMES; A PARADIGM SHIFT IN NGS-BASED GENOTYPING

Bode A. Olukolu[1] and G. Craig Yencho[2]

[1] Assistant Professor, Department of Entomology and Plant Pathology, University of Tennessee, Knoxville, TN USA 37996 (bolukolu@utk.edu); [2] Professor, Department of Horticultural Science, NC State University, Raleigh, NC USA 27695.

Next Generation Sequencing (NGS) is an extensively used tool for massive parallel sequencing of genomes. However, several applications such as NGS-based genotyping still suffer from pitfalls that limit their accuracy and utility. To mitigate these pitfalls, we have developed a significant advancement in short-read NGS library preparation. Presented here is an inexpensive quantitative reduced representation sequencing (qRRS) approach for dosage-sensitive genotyping and quantitative strain-level metagenome/microbiome profiling. The scalable, ligation-free and double-stranded DNA-protection assay eliminates off-target annealing temperature-dependent hybridization. This is achieved by using single-stranded barcoded adapters for isothermal strand displacement of double-stranded DNA templates with restriction site overhangs as the only priming site. As much as 9,216 samples can be multiplexed. Novel features in this protocol include a paradigm shift in adapter design that prevents chimeric reads and barcode swapping, a flow cell cluster enhancer that generates about 50% more yields, and consistent high-quality base calling scores. The library preparation workflow is optimized for ease-of-use and can be completed in one to two days. To accommodate these novel features during data pre-processing and downstream analytics, we have developed bioinformatic and analytical pipelines for empirical-based NGS data quality filtering (ngsComposer), haplotype-based variant calling and filtering (GBSapp), and quantitative metagenomic alignment and taxonomic exact matching (Qmatey). The qRRS approach establishes new standards in high-fidelity quantitative genotyping, minimizes missing data and allelic dropout, and makes functional microbiome studies more accessible. Compared to 16S amplicon sequencing, which uses a single gene for microbiome profiling, qRRS provides multiple genome-wide sequences for strain-level taxonomic delineation and quantification. We are now exploring its utility as a diagnostic tool for scoring multiple diseases (and disease complexes) based on titer levels of pathogens in a single assay. We envision that the enhanced quality and quantity of qRRS-derived markers will improve genomic-assisted breeding efforts.