

## ***Aspergillus flavus* Pangenome to Capture the Diversity in a Single Reference Genome**

**S. GANGURDE\***, H. WANG, R. KEMERAIT, B. DUTTA, Department of Plant Pathology, University of Georgia, Tifton, GA; J. FOUNTAIN, Department of Biochemistry, Molecular Biology, Entomology and Plant Pathology, Mississippi State University; P. BAJAJ, M. PANDEY, International Crops Research Institute for Semi-Arid Tropics (ICRISAT), Hyderabad, India; H. ABBAS, USDA-ARS, Biological Control of Pests Research Unit, Stoneville, MS; P.K. CHANG, USDA-ARS, Southern Regional Research Center, New Orleans, LA; B. GUO, USDA-ARS, Crop Genetics and Breeding Research Unit, Tifton, GA.

There were several chromosome-level reference genomes published recently for *Aspergillus flavus*, a significant advancement for genomic studies. However, a single reference genome creates reference bias in the genomic analysis, such as a 310Kb insertion was reported for isolate AF13 and which was missing in isolate NRRL3357. A regulator gene *atfC* has been identified within this insertion of AF13, a high aflatoxin producer. To tackle the reference bias a pangenome framework can be developed, which could include the variation and diversity of all the genes from one species. An *A. flavus* pangenome can be used as a reference genome to discover novel aflatoxin regulator genes. Here we sequenced a total of 221 isolates of diverse origin, including 98 isolates from various parts of corn plants in Mississippi and 161 isolates (123 *A. flavus* used for this pangenome and 38 *A. parasiticus* excluded) from various soils of different cropping systems in Georgia. In addition, we also included 125 isolates from public domain. Finally, the whole genome sequencing data for 346 isolates was used to develop scaffold level assemblies. The genome size of these isolates was in a range of 36.5Mb-43.5Mb with an average genome size of 37.37 Mb. A total of 1.8 million genome wide SNP variants were discovered across the isolates. Population structure analysis identified six sub-populations, indicating significant genome level diversity among the isolates. The isolates from corn clusters separately with slight admixture, indicating unique genome composition of isolates associated with different parts of corn plants. Furthermore, 346 assemblies were used to identify the core and the accessory genomes and a pangenome framework with variations from all isolates have been developed. We are investigating the variation of aflatoxin producing gene clusters among the isolates and in comparison, to *A. parasiticus*. The mutations associated with increased or reduced aflatoxin biosynthesis will also be investigated. This pangenome of *A. flavus* will be used for Pan-GWAS analysis to demonstrate the potential of pangenome to identify the genes associated with aflatoxin and secondary metabolites production.