

## **Assembly of *de novo* Genome Sequence Increases Candidate Gene Discovery: A Case of NC94022 and TSWV Resistance**

**J. CLEVINGER\***, Center for Plant Genomics and Sustainable Agriculture, Hudson Alpha Institute for Biotechnology, Huntsville, AL 35802; B. SCHEFFLER, S. SIMPSON, C. YOUNGBLOOD, USDA-ARS Genomics, Bioinformatics Research Unit, Stoneville, MS; H. WANG, Department of Plant Pathology, The University of Georgia and USDA-ARS, Tifton, GA; A. CULBREATH, Department of Plant Pathology, The University Of Georgia; X. WANG and C. ZHAO, Shandong Academy of Agricultural Sciences, Jinan, China; and B. GUO, USDA-ARS, Tifton, GA 31793.

To facilitate fine mapping and candidate gene discovery for the strong Tomato Spotted Wilt Virus (TSWV) resistance from the *Arachis hypogaea* var *hirsuta* – derived line NC94022, a *de novo* genome assembly was constructed using Pacific Bioscience long reads. A final assembly of 2.5 Gb with a contig N50 of 1.8 Mb was scaffolded into 20 pseudomolecules. Comparative analysis with Tifrunner shows low single nucleotide polymorphism (SNP) diversity, further supporting extremely low SNP diversity within *A. hypogaea*. Of note are significant insertions and deletions that contribute differential gene content. Targeted analysis of the mapped region on A01 controlling TSWV resistance identified only 1 SNP per 22 kb, and only 7 SNPs in the estimated 200 kb region controlling resistance. An insertion was identified in the region resulting from a duplication of two genes. The comparative genome analysis reveals how *de novo* genome assembly delivers the toolset needed to identify candidate functional variation, and even genes, for marker development rapidly and effectively. We recommend that any mapping experiment include *de novo* assembly of parental genomes for effective marker development and gene discovery.