

B-176

Assembly and annotation of a draft genome sequence for *Glycine latifolia* - a perennial wild relative of soybean

Qiong Liu*, Department of Crop Sciences, University of Illinois, Illinois, USA

Leslie Domier, Department of Plant Pathology, USDA-ARS, Illinois, USA

Glycine latifolia (Benth.) Newell & Hymowitz, is one of the 26 wild perennial relatives of soybean (*Glycine max* (L.) [Merr]) in the genus *Glycine* subgenus *Glycine*. Because selected accessions of perennial *Glycines* species are capable of outperforming soybean under multiple abiotic and biotic stresses, there has been a growing interest in utilizing them as genetic sources for soybean improvement. To better understand the genomic characteristics of a *Glycines* species, we conducted whole genome assembly and annotation of *G. latifolia* using exclusively Chromium-linked reads (10x Genomics) sequenced from a single barcoded library. The Supernova assembler yielded a final scaffold assembly of 989.6 Mb (scaffold N50=853.6 kb) representing 87.6% of the estimated 1.13-Gb genome. With the assistance of genetic maps of *G. latifolia* and the soybean genome sequence, the scaffold assembly was further organized into 20 chromosome-scale pseudomolecules and 41,231 unplaced scaffolds (≥ 1 kb). Comparison of chromosome sequences between *G. latifolia* and the soybean genome confirmed and complemented the discovery of interchromosomal rearrangements from previous studies. We subsequently performed genome annotation following the MAKER annotation pipeline based on the integrated evidence of *G. latifolia* transcriptome, *ab initio* prediction, and soybean protein sequences. Using the predicted protein sequences, regions of extensive synteny were observed in 12 chromosomes between *G. latifolia* and soybean. Comparative analysis was conducted to detect orthologous gene clusters between the *G. latifolia* genome and several sequenced legume genomes. The results from multiple assembly assessments, including transcriptome read alignment (>90%) and BUSCO (92.40%), suggested a high quality genome assembly with excellent completeness. This draft genome sequence of *G. latifolia* will serve as a valuable source of genetic information for identification of new alleles and genes for soybean improvement.