# ASSEMBLY AND ANNOTATION OF THE GREEN ASH GENOME

Matthew Huff [1*], Josiah Seaman[2,3], Di Wu[4], Tetyana Zhebentyayeva[4], Laura Kelly[2,3], Nurul Faridi[5,6], C. Dana Nelson[6,7], Endymion Cooper[2], Enrico Bonello[8], Jennifer Koch[9], Jeanne Romero-Severson[10], John E. Carlson[4], Richard Buggs[2,3] and Margaret Staton[1]

[1] Department of Entomology and Plant Pathology, University of Tennessee, Knoxville, Tennessee, USA
[2] School of Biological and Chemical Sciences, Queen Mary University of London, London, UK
[3] Royal Botanic Gardens, Kew, Richmond, Surrey, UK
[4] Department of Ecosystem Science and Management, Pennsylvania State University, University Park, Pennsylvania, USA.
[5] Department of Ecosystem Science and Management, Texas A&M University, College Station, Texas, USA.
[6] USDA Forest Service, Southern Research Station, Saucier, Mississippi, USA
[7] Forest Health Research and Education Center, University of Kentucky, Lexington, KY, USA
[8] Department of Plant Pathology, The Ohio State University, Columbus, Ohio, USA
[9] United States Department of Agriculture, Forest Service, Northern Research Station, Delaware, Ohio, USA
[10] Department of Biological Sciences, Notre Dame University, 46556 Notre Dame, Indiana, USA

Once the most widely distributed ash tree in North America, the green ash (*Fraxinus pennsylvanica*) has faced high mortality as a result of the non-native invasive emerald ash borer (EAB; *Agrilus planipennis*). A small percentage of native green ash trees that remain healthy in long-infested areas, termed "lingering ash," display partial resistance to the insect, indicating that breeding and propagating populations with higher resistance to the pest may be possible. Genomic resources could assist green ash resistance breeding programs by enabling identification of genetic markers associated with resistance and other important traits. However, at present only scaffold-level genome assemblies, without gene or trait information, are available. As a first step toward providing information needed by tree breeders, we developed a 757 Mbp chromosome-scale assembly scaffolded with a newly expanded genetic linkage map containing over 4,000 SNP markers. Gene annotation of the assembled genome sequence identified 35,470 protein coding genes. Synteny and base-pair substitution analysis confirmed the presence of the previously reported Oleaceae family whole genome duplication. Interestingly, the chromosomes are only moderately rearranged since the duplication event and residual syntenic blocks are identified both within the green ash genome and between green ash and olive. In addition, we demonstrate further utility of the new assembly through reference-guided scaffolding of publicly available genomes from 28 other *Fraxinus* species and subspecies. These resources and analyses provide a new opportunity to characterize EAB response among species spanning from resistant to susceptible in phenotype and should benefit EAB-resistance breeding programs and ash restoration efforts.