

A PHASED, CHROMOSOME-SCALE GENOME ASSEMBLY FOR SOUTHERN LIVE OAK, *QUERCUS VIRGINIANA*

Laramie Aközbek, Zachary Meharg, Claudia Ann Rutland, Rijan Dhakal, Tosin Akinspe, Jillian Abendroth, Nicholas Gladstone, Sejal Patel, Zahida Pervaiz, Giovanni Rossi, F. Alex Feltus, David J. Clarke IV, John Lovell, Jeremy Schmutz, Jenell Weber, Lori Beth Boston, Jane Grimwood, Haley Hale, Sarah Carey, Leslie Goertzen, and Alex Harkess

PURPOSE AND SCOPE

Like many members of *Quercus* subgenus *Quercus*, the southern live oak (*Quercus virginiana*) often remains genetically distinct despite living in sympatry and experiencing repeated interspecific hybridization (Cavender-Bares and Pahlich 2009). A comprehensive characterization of the oak syngameon requires the

Laramie Aközbek, Department of Crop, Soil, and Environmental Sciences, Auburn University, Auburn, AL 36849, USA and HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA

Zachary Meharg, Department of Crop, Soil, and Environmental Sciences, Auburn University, Auburn, AL 36849, USA and HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA

Claudia Ann Rutland, Department of Crop, Soil, and Environmental Sciences Auburn University, Auburn, AL 36849, USA

Rijan Dhakal, Department of Biological Sciences, The University of Alabama in Huntsville, Huntsville, AL, 35899, USA and Bioinformatician, Donald Danforth Plant Science Center, Olivette, MO 63132, USA

Tosin Akinspe, Department of Crop, Soil, and Environmental Sciences, Auburn University, Auburn, AL 36849, USA

Jillian Abendroth, Department of Crop, Soil, and Environmental Sciences, Auburn University, Auburn, AL 36849, USA

Nicholas Gladstone, Department of Crop, Soil, and Environmental Sciences, Auburn University, Auburn, AL 36849, USA

Sejal Patel, Department of Crop, Soil, and Environmental Sciences, Auburn University, Auburn, AL 36849, USA

Zahida Pervaiz, Department of Crop, Soil, and Environmental Sciences, Auburn University, Auburn, AL 36849, USA

Giovanni Rossi, Department of Crop, Soil, and Environmental Sciences, Auburn University, Auburn, AL 36849, USA

F. Alex Feltus, Department of Genetics and Biochemistry, Clemson University, Clemson, SC 29634 and Praxis AI LLC, Sacramento, CA 95816, USA

David J. Clarke IV, Praxis AI LLC, Sacramento, CA 95816, USA

John Lovell, Genome Sequencing Center, HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA

Jeremy Schmutz, Genome Sequencing Center, HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA

Jenell Weber, Genome Sequencing Center, HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA

Lori Beth Boston, Genome Sequencing Center, HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA

Jane Grimwood, Genome Sequencing Center, HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA

Haley Hale, Center for Plant Science and Sustainable Agriculture, HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA

Sarah Carey, Center for Plant Science and Sustainable Agriculture, HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA

Leslie Goertzen, Department of Biological Sciences, Auburn University, Auburn, AL 36849, USA

Alex Harkess, Center for Plant Science and Sustainable Agriculture, HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA

synthesis of population genetic and genomic approaches. High-quality reference genomes are essential to the goal of understanding this phenomenon, allowing us to more precisely describe the landscape of recombination and genetic variation in target species. As part of the American Campus Tree Genome Project, the goal was to develop a fully phased, chromosome-scale genome assembly for *Q. virginiana* to aid in efforts to improve and understand oak genetics.

METHODS AND APPROACH

Tissue samples for DNA and RNA were collected from a young clone of Auburn University's "Toomer's Oak" (*Q. virginiana*). The genome size, ploidy, and heterozygosity were estimated with Illumina short reads. First, the mitochondrial and chloroplast genomes were constructed from long reads with *oatk*. Then, the nuclear genome was assembled with *hifiasm* utilizing PacBio Hifi reads and scaffolded utilizing Dovetail Omni-C short-reads with *YAHS*. RNA was extracted from young and senescing leaves as well as roots, and subsequently used to annotate the genome with *Braker3*. Repetitive elements were characterized with *EDTA*, and comparative genomic analyses with other *Quercus* and Fagales members were conducted using the *GENESPACE* package.

FINDINGS AND IMPLICATIONS

The resulting diploid ($2n = 24$) assembly was highly contiguous and complete. The final haplotype assemblies included HAP1 with 777.1 Mb in 175 total contigs (contig N50 = 48.1 Mb, L50 = 7) and HAP2 with 769.9 Mb in 132 contigs (contig N50 = 53 Mb, L50 = 7). The k-mer completeness of the combined assemblies was 97.5 percent (QV 42.9). Repeats represent approximately 60 percent of the sequence in both haplotypes. Approximately 30,500 gene models were annotated in both haplotypes, consistent with previous annotations of species in this genus.

Initial comparative analyses suggest an extraordinary level of synteny between this species and other members of subgenus *Quercus*. Extensive chromosomal rearrangements are common among plants and these rearrangements can impact hybridization between related species. This unusual degree of conserved gene order may provide insights into the biology of the syngameon as more oak genomes are released. Additionally, alongside other references, the *Q. virginiana* genome will contribute to further research into all aspects of oak biology, such as longevity, biotic and abiotic environmental adaptations, disease resistance, and more.

LITERATURE CITED

Cavender-Bares, J.; Pahlich, A. 2009. Molecular, morphological, and ecological niche differentiation of sympatric sister oak species, *Quercus virginiana* and *Q. geminata* (Fagaceae). *American Journal of Botany*. 96(9): 1690–1702. <https://doi.org/10.3732/ajb.0800315>.

ACKNOWLEDGMENTS

The authors thank the organizers and contributors to Tiger Giving Day, and the NSF IOS-PGRP CAREER #2239530 to AH, which provided funding for this project.