

Sequencing hundreds of nuclear genes in the sunflower family now possible

February 20 2014

Advances in DNA sequencing technologies have enormous potential for the plant sciences. With genome-scale data sets obtained from these new technologies, researchers are able to greatly improve our understanding of evolutionary relationships, which are key to applications including plant breeding and physiology.

Studies of evolutionary (or phylogenetic) relationships among different plant species have traditionally relied on analyses of a limited number of genes, mostly from the chloroplast genome. Such studies often fail to fully or accurately resolve [phylogenetic relationships](#), given the limited amount of data used.

New methods of DNA sequencing have made it possible for researchers to sequence hundreds to thousands of specific nuclear genes, greatly facilitating studies of phylogenetic relationships. However, despite the great potential of this approach, termed "target sequence capture," few researchers have developed protocols to sequence numerous nuclear genes for plant [phylogenetic studies](#).

Researchers at the University of Memphis, the Smithsonian Institution, the University of Georgia, and other institutions have designed an efficient approach for sequencing hundreds of nuclear genes across members of the Compositae (the sunflower family). The Compositae are one of the largest families of flowering plants, containing around 25,000 species and numerous economically important crop plants, such as lettuce, sunflower, and artichoke, as well as numerous ornamentals.

The new protocol (available for free viewing in the February issue of *Applications in Plant Sciences*) will allow researchers to better-resolve phylogenetic relationships at both deep and shallow levels within the family, providing an excellent framework for addressing evolutionary questions about the family. Previous phylogenetic studies of the family, based on up to 10 [chloroplast genes](#), had failed to resolve certain key relationships, limiting inferences of morphological evolution.

According to Jennifer Mandel, assistant professor in the Department of Biological Sciences at University of Memphis and lead author of the paper, the new approach is an improvement on traditional, PCR-based sequencing strategies, which have generally focused on chloroplast genes or a handful of nuclear genes. "Our method samples the genome much more widely, while avoiding the repetitive regions that make many plant genomes so difficult to assemble," says Mandel.

The protocol employs custom-designed probes that can hybridize with and "capture" 1061 [nuclear genes](#) from DNA samples of sunflower species. The captured genes can then be sequenced on the Illumina HiSeq or a similar next-generation sequencing platform, allowing tremendous amounts of data to be recovered for phylogenetic analysis.

The researchers also developed a bioinformatic and phylogenetic workflow for processing and analyzing the resulting sequence data. The workflow assembles the genes from the millions of reads generated from the sequencing instrument and then assesses all of the recovered genes for orthology (i.e., for their ability to reflect speciation events and, therefore, to accurately reconstruct phylogenetic relationships). The genes that pass the orthology test are then used for large-scale phylogenetic analyses.

The researchers tested the efficacy of the probes and overall workflow using 14 species from the family (and one from its closest relative,

Calyceraceae). The species selected span the phylogenetic breadth of the family, allowing the researchers to assess the utility of the method at broad taxonomic levels. Several closely related species (from the tribe Heliantheae) were also included to assess the usefulness of the method for shallow phylogenetic studies within the Compositae.

The researchers were able to successfully recover a large portion of the 1061 target genes across all the species included, and around 700 of these genes were determined to be orthologous and thus suitable for [phylogenetic analysis](#). Using these orthologous [genes](#), they were able to generate well-resolved phylogenetic trees consistent with known relationships in the family, demonstrating the successfulness of this approach for phylogenetic studies of the Compositae.

Although the probe set was developed specifically for research on the sunflower family, the researchers note that the overall workflow can be applied to any taxonomic group of interest. Therefore, this protocol could serve as a model for phylogenetic investigations of other major plant groups, as well as an excellent tool for studies of the Compositae.

"Novel probes can be designed as long as transcriptomic data exists or can be gathered for the taxa of interest," says Mandel.

More information: Jennifer R. Mandel, Rebecca B Dikow, Vicki A. Funk, Rishi R. Masalia, S. Evan Staton, Alex Kozik, Richard W. Michelmore, Loren H. Rieseberg, and John M. Burke. 2014. A target enrichment method for gathering phylogenetic information from hundreds of loci: An example from the Compositae. *Applications in Plant Sciences* 2(2): 1300085. [DOI: 10.3732/apps.1300085](https://doi.org/10.3732/apps.1300085)

Provided by American Journal of Botany

Citation: Sequencing hundreds of nuclear genes in the sunflower family now possible (2014, February 20) retrieved 10 May 2024 from <https://phys.org/news/2014-02-sequencing-hundreds-nuclear-genes-sunflower.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.