

## **Two Shifts in Evolutionary Lability Explain Many Independent Origins of Nitrogen-Fixing Nodulation Symbiosis in A Single Clade of The Plant Tree of Life**

Heather R. Kates<sup>1\*</sup> (hkates@ufl.edu), Ryan A. Folk<sup>2</sup>, Pamela S. Soltis<sup>1,3,4</sup>, Robert P. Guralnick<sup>1,5</sup>, Brian O'Meara<sup>6</sup>, Raphael LaFrance<sup>1</sup>, **Matias Kirst**<sup>7</sup>, Douglas E. Soltis<sup>1,3,4,5</sup>

<sup>1</sup>Florida Museum of Natural History, University of Florida, Gainesville, FL, United States;

<sup>2</sup>Department of Biological Sciences, Mississippi State University, Mississippi State, MS, United States; <sup>3</sup>Department of Biology, University of Florida, Gainesville, FL, United States;

<sup>4</sup>Biodiversity Institute, University of Florida, Gainesville, FL, United States; <sup>5</sup>Genetics Institute, University of Florida, Gainesville, FL, United States; <sup>6</sup>Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN 37996-1610, USA; <sup>7</sup>Forest Resources and Conservation, University of Florida, Gainesville, FL, United States

[www.nitfix.org](http://www.nitfix.org)

**Project Goals: The collaborative project "Phylogenomic discovery and engineering of nitrogen fixation into the bioenergy woody crop poplar" is focused on identifying the genomic novelties that enable the symbiotic relationship between nodulating plants and N-fixing bacteria to support genetically engineering this capability into bioenergy crops. The first aim of this project is a comparative phylogenomic study of the nodulating clade to uncover the genomic novelties that were required for the evolution of these root nodule symbioses. A massively improved phylogenetic framework for the N-fixing clade that includes genetic data for nodulation genes for all species in the phylogeny will provide a robust, revised understanding of the exact ancestral origin of nodulation and the evolution of the predisposition to nodulate and inform a series of experiments comparing close relatives that nodulate or do not nodulate to identify the genes underlying nodulation. To achieve these goals we have completed a cutting edge phylogenetic study with an unprecedented sampling effort and used this to identify the origins of the symbiotic relationship between nodulating plants and N-fixing bacteria as illustrated by the project described below.**

Symbiotic nitrogen fixation via root nodulation (SNFN) is a complex trait that requires coordinated control of multiple plant host and bacterial symbiont genes. Pinpointing the evolutionary origins of SNFN is critical for understanding its genetic basis but is complicated by data limitations and intermittent presence in a single species-rich clade of flowering plants, the nitrogen-fixing clade (NFC). Here we use the largest purpose-built phylogeny for any lineage, along with an enhanced trait database, to reconstruct the evolution of SNFN. We show that shifts among heterogeneous evolutionary rates can explain how a trait can arise many times across a large phylogeny by identifying the evolutionary pathway to SNFN gain. Our analyses suggest a two-step process where an ancestral deep precursor state gives rise to a more labile state from which SNFN was easily gained at certain points in the NFC. Our reconstruction of ancestral

states illustrates how this two-step pathway could have led to 15 independent gains of SNFN, in contrast to single-gain and losses hypotheses. SNFN may be an example of multi-level convergent evolution, thus requiring scaling up of genome-phenome mapping to fully elucidate mechanisms enabling SNFN.

**Funding Statement:** This research was supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-1916632.