

Title: Understanding Nitrogen Use Efficiency and Oilseed Traits in Camelina by High Resolution Genome Sequencing and Whole-genome Resequencing

Authors: Charlemagne Lim^{1*}(charlemagnealexalim@montana.edu), Samuel Decker^{1*}(samuel.decker@student.montana.edu), Huang Li¹, John Lovell², Paul Grabowski², Jeremy Schmutz², Jennifer Lachowiec¹, Chengci Chen¹, **Chaofu Lu**¹

Institutions: ¹Montana State University, Bozeman; ²HudsonAlpha Institute for Biotechnology, Huntsville, AL.

Website URL: <https://www.montana.edu/econproject/index.html>

Project Goals: Camelina is a Brassica oilseed crop that has great potential to become a sustainable source of bioenergy in the US. However, the low nitrogen use efficiency and the low seed and oil yield compared to other major oilseed crops hinder this potential. The goal of this project is to decipher the genetic and physiological mechanisms that determine the nitrogen use efficiency and oilseed yield during the most critical processes of the camelina life cycle: 1) how camelina, in partnership with soil microbes, maximizes its ability to absorb and assimilate nitrogen into vegetative biomass; and 2) upon the transition to reproductive growth, how nitrogen is efficiently remobilized from senescing tissues (leaves and silicles) into sinks (seeds) to optimize yield potential by increasing seed size and enhancing oil synthesis.

Abstract Text:

Camelina is an emerging oilseed crop that has great potential as a renewable, high quality, and sustainable source of bioenergy in the US. However, this great potential is offset by camelina's low nitrogen use efficiency and low seed and oil yield compared to other oilseed crops. Development of camelina varieties with superior seed and oil yield characteristics requires careful examination and the decryption of the underlying genetic and physiological mechanisms critical in the camelina life cycle that govern the foraging and uptake of nutrients from the soil, source-sink relationships and resource allocation, reproduction, seed formation, and oil synthesis. This can be facilitated by new technological advancements in biotechnology in conjunction with modern breeding techniques. To understand the natural variation within camelina species, we determined the population structure of a worldwide collection of 222 *Camelina sativa* accessions using 161,301 single nucleotide polymorphisms (SNPs) derived from whole-genome resequencing. Additionally, when coupled with genome-wide association studies (GWAS) and linkage mapping using a recombinant inbred line population, QTL and candidate genes that control field-evaluated agronomic traits related to oilseed production such as seed size, oil

content, fatty acid composition, and flowering time were identified. To further locate these QTLs and candidate genes with high precision, we have assembled a high-quality camelina genome through Pac-bio HiFi genome sequencing coupled with Hi-C genome scaffolding. The sequenced camelina line (Suneson) has been used in our mapping studies and will provide more accurate genetic markers. Field trials under low and high nitrogen soils were conducted in North East and Central Montana, and in Eastern Washington, US in 2021 (and to be repeated in 2022) to generate a new set of phenotypic data from the accessions, which were used to determine the relationships among morpho-agronomic traits and for the GWAS to identify QTLs and candidate genes for increased nitrogen use efficiency and oilseed traits. Identification of QTL and genes together with their genetic markers will greatly facilitate breeding efforts to develop camelina varieties with high nitrogen use efficiency, high seed yield and oil quality, and are adaptable to the US northwest environments.

References/Publications:

1. Li H, Hu X, Lovell JT, Grabowski PP, Mamidi S, Chen C, Amirebrahimi M, Kahanda I, Mumey B, Barry K, Kudrna D, Schmutz J, Lachowiec J, & Lu C. Genetic dissection of natural variation in oilseed traits of camelina by whole-genome resequencing and QTL mapping. *Plant Genome*. 2021;14:e20110. <https://doi.org/10.1002/tpg2.20110>

Funding Statement: This research is supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomic Science Program grant no. DE-SC0021369.