

114. Comparative Genomics in Support of CAM Engineering

Xiaohan Yang^{1*} (yangx@ornl.gov), Hengfu Yin¹, Susanna F. Boxall², James Hartwell², David Weston¹, Paul E Abraham¹, Timothy J. Tschaplinski¹, Sara Jawdy¹, Priya Ranjan¹, Gerald A. Tuskan¹, Robert L. Hettich¹, Stan D. Wullschleger¹, Bernard Wone³, Won Cheol Yim³, Jungmin Ha³, **John C. Cushman**³, Anne M. Borland^{1,4}, Karen Schlauch³, Rachel Brenchley², Richard Gregory², Neil Hall², Tanja Magoc⁵, Steven L. Salzberg⁶, Stephen M. Gross⁷, Axel Visel⁷, Zhong Wang⁷, June Simpson⁸

¹Oak Ridge National Laboratory, Oak Ridge, TN; ²University of Liverpool, Liverpool; ³University of Nevada-Reno, Reno, NV; ⁴Newcastle University, Newcastle upon Tyne; ⁵University of Maryland, College Park, MD; ⁶Johh Hopkins University School of Medicine, Baltimore, MD; ⁷Lawrence Berkeley National Laboratory, Berkeley, CA; ⁸CINVESTAV, Irapuato, Guanajuato.

<http://cambiodesign.org/>

Project Goals: The long-term goal of this project is to enhance the water use efficiency (WUE) and adaptability to hotter, drier climates of species that normally perform C₃ photosynthesis by introducing Crassulacean Acid Metabolism (CAM). To achieve this long-term goal, four major objectives are being pursued: 1) define the genetic basis of key CAM modules; 2) characterize the regulation of ‘carboxylation’, ‘decarboxylation’, and ‘inverse stomatal control’ modules of CAM; 3) deploy advanced genome engineering technologies to enable stacking of a large number of transgenes into a single genomic locus to transfer fully functional ‘carboxylation’ and ‘decarboxylation’ modules from CAM species to C₃ species that can accommodate overnight malic acid storage in the vacuole; and 4) analyze the effects of these transgenic modules on ‘stomatal control’, CO₂ assimilation and transpiration rates, biomass yield, and WUE in *Arabidopsis* and *Populus*. CAM biodesign research is multidisciplinary by design and multi-institutional in composition.

Improved crop water-use efficiency (WUE) is critical for the long-term sustainability of agricultural production systems in the face of predicted future warmer and drier climates. Crassulacean acid metabolism (CAM) is a specialized mode of photosynthesis that enhances WUE through an inverse day/night pattern of stomatal closure/opening and improves photosynthetic efficiency by concentrating CO₂ around RUBISCO. CAM has evolved multiple times from C₃ photosynthesis and ~6.5% of higher plant species in more than 35 families have acquired CAM via parallel or convergent evolution. There are three fundamental questions to be answered to understand the molecular basis and evolutionary mechanism of CAM: 1) how many genes are needed for the C₃-to-CAM transition? 2) what are these genes? and 3) when did they evolve? To address these questions, we used an approach integrating co-expression gene network and protein sequence evolution, with a focus on the difference between CAM and non-CAM species as well as the conservation among different CAM species. Our network analysis based on *Agave* RNA-seq data identified co-expression gene modules associated with the CAM pathway. Through comparative genomics analysis of 15 species selected from diverse lineages, including CAM (e.g., *Agave*), C₃ (e.g., *Arabidopsis*, *Oryza*, *Populus*), C₄ (e.g., *Setaria*, *Sorghum*, *Zea*), and non-vascular plants (e.g., *Physcomitrella*, *Selaginella*), we found that the core components of CAM machinery have an ancient origin traceable to non-vascular plant lineages, and regulatory proteins were essential to the C₃-to-C₄ transition. Furthermore, our comparison of multiple CAM plants (i.e., *Agave*, *Kalanchoe*, *Mesembryanthemum*) revealed that parallel or

convergent evolution resulted in regulatory machinery that transitioned from the C₃ to CAM pathways. This research establishes a framework for CAM comparative genomics studies and provides new knowledge to inform genetic improvement in WUE and photosynthetic efficiency in crop plants under water-limiting conditions using synthetic biology approaches.

***Acknowledgement:** This research was supported by the Department of Energy, Office of Science, Genomic Science Program under Award Number DE-SC0008834. Oak Ridge National Laboratory is managed by UT-Battelle, LLC for the U.S. Department of Energy under Contract Number DE-AC05-00OR22725.*