

## Kinetic modeling of the phenylpropanoid pathway in Arabidopsis

Longyun Guo<sup>1\*</sup> (guo165@purdue.edu), Peng Wang<sup>1</sup>, Rohit Jaini<sup>2</sup>, Natalia Dudareva<sup>1</sup>, Clint Chapple<sup>1</sup> and **John Morgan**<sup>1,2</sup>

<sup>1</sup>Department of Biochemistry, Purdue University, West Lafayette, IN; <sup>2</sup>School of Chemical Engineering, Purdue University, West Lafayette, IN-47907

**Project Goals:** The project aims to generate a kinetic model of lignin biosynthesis in Arabidopsis to guide rational design for biofuel production. Lignin biosynthesis requires 11 enzyme families functioning together to produce 3 major monolignols from phenylalanine. The metabolic complexity makes it difficult to easily anticipate outcomes of metabolic engineering. A kinetic model is thus proposed to help both understanding and manipulation of this pathway. We are iteratively developing the model with data from *in vivo* substrate feeding to Arabidopsis stems. In parallel, we are also working on introducing an alternative pathway to 2-phenylethanol into Arabidopsis. The final goal is to rewire the fluxes around and downstream of phenylalanine to design an efficient biofuel-producing plant.

Lignocellulosic biomass in plant is an important source for bio-ethanol production, however, its efficient breakdown is limited by the cross-linking property of the lignin polymer in the secondary cell wall. To solve the problem, it is promising to develop transgenic plants with reduced amount of lignin or altered lignin composition. Lignin is mainly derived from *p*-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol. These alcohols are synthesized in the cytosol from the common precursor phenylalanine. During these metabolic conversions, there are cases where multiple substrates compete for a common enzyme, and where single intermediates are funneled through multiple branch points. This complex metabolic network makes it challenging to rationally propose a genetic manipulation strategy for desired product yield, due to the lack of a global understanding of the pathway. An integrative experimental-mathematical framework was thus proposed to generate an *in vivo* kinetic model for both basic understanding and to serve as an *in silico* platform for manipulation outcome prediction. The Arabidopsis primary stem was selected as the experimental system for modeling lignin formation. Different concentrations of <sup>13</sup>C<sub>6</sub>-ring labeled phenylalanine were fed to excised 5-week-old stems to obtain both the amount and isotopic enrichment of downstream intermediates at multiple time points with LC-MS/MS. Meanwhile averaged lignin deposition rate was estimated from lignin content over development. Since a complete kinetic model is computationally intensive to parameterize, we decided to divide the whole pathway into three modules to train smaller models respectively. A base kinetic model for the first module was initially constructed using Michaelis-Menten kinetics. For the following model refining step, 36 possible metabolite-enzyme regulatory interactions were then systematically explored. Evidence of their existence(s) *in vivo* was examined with training datasets (0.1, 1 & 3 mM treatments) using Akaike's Information Criteria. The best performance model was then validated with an independent

dataset (0.3 mM treatment). With this workflow, we identified several previously unknown putative metabolite-enzyme interactions, and the current model can capture pathway dynamics over a wide range of feeding treatments. Base models for the other two modules have also been constructed, and similar workflow is planned to apply for model refining as well. The final combined kinetic model can be used to explore *in vivo* metabolic behaviors under different conditions.

*This research is supported by the award DE-SC0008628 from the Office of Biological and Environmental Research in the US Department of Energy.*