Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

Flux modeling for monolignol biosynthesis

Jack P Wang[1,2], Megan L Matthews[3], Punith P Naik[4,5], Cranos M Williams[3], Joel J Ducoste[5], Ronald R Sederoff[2] and Vincent L Chiang[1,2]

The pathway of monolignol biosynthesis involves many components interacting in a metabolic grid to regulate the supply and ratios of monolignols for lignification. The complexity of the pathway challenges any intuitive prediction of the output without mathematical modeling. Several models have been presented to quantify the metabolic flux for monolignol biosynthesis and the regulation of lignin content, composition, and structure in plant cell walls. Constraint-based models using data from transgenic plants were formulated to describe steady-state flux distribution in the pathway. Kinetic-based models using enzyme reaction and inhibition constants were developed to predict flux dynamics for monolignol biosynthesis in wood-forming cells. This review summarizes the recent progress in flux modeling and its application to lignin engineering for improved plant development and utilization.

Addresses
[1] State Key Laboratory of Tree Genetics and Breeding, Northeast Forestry University, Harbin 150040, China
[2] Forest Biotechnology Group, Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC 27695, United States
[3] Electrical and Computer Engineering, North Carolina State University, Raleigh, NC 27695, United States
[4] Freddie Mac, McLean, VA 22102, United States
[5] Civil, Construction and Environmental Engineering, North Carolina State University, Raleigh, NC 27695, United States

Corresponding author: Chiang, Vincent L (vchiang@ncsu.edu)

Current Opinion in Biotechnology 2019, **56**:187–192

This review comes from a themed issue on **Plant biotechnology**

Edited by **Wout Boerjan** and **John Ralph**

For a complete overview see the Issue and the Editorial

Available online 18th December 2018

https://doi.org/10.1016/j.copbio.2018.12.003

0958-1669/© 2018 Elsevier Ltd. All rights reserved.

Lignin is a complex phenolic polymer in the cell wall of secondary-thickened plant cells and one of the most abundant terrestrial biopolymers [1]. Lignin is crucial for not only vascular transport, but also mechanical support, and provides defense against pests and pathogens [2]. Lignin is formed by the oxidative polymerization of one or more types of major monolignols including sinapyl, coniferyl, and 4-coumaryl alcohols that are incorporated into the polymer resulting in syringyl (S), guaiacyl (G), and *p*-hydroxyphenyl (H) units, respectively [3,4].

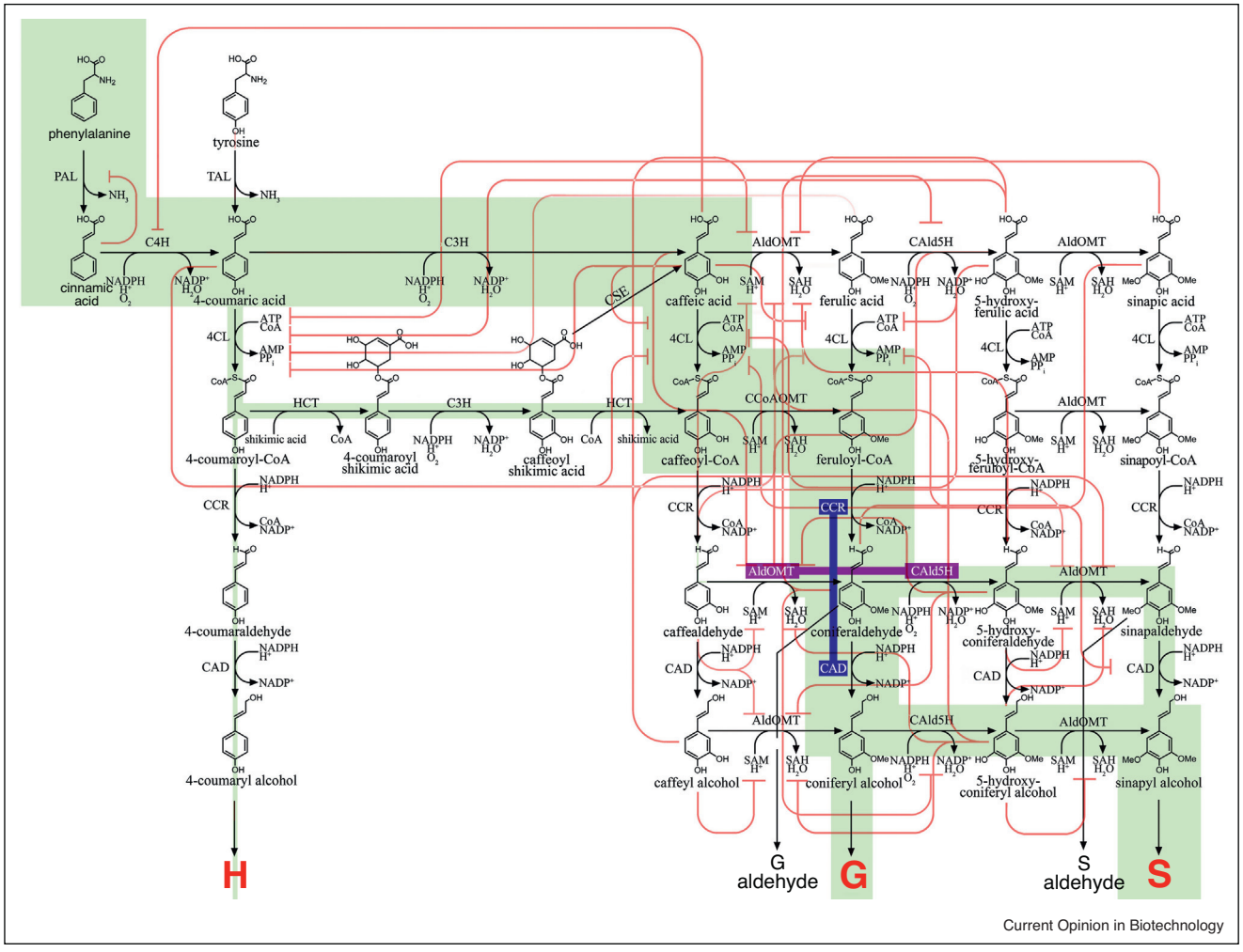
The biosynthetic pathway (Figure 1) for lignin has been studied extensively because of its fundamental role in the regulation of plant secondary cell wall formation and the consequent determination of wood chemical and physical properties [5,6••,7–9]. Lignin content, composition, and structure are predominantly controlled by the metabolic flux that regulates the supply and ratio of monolignols for lignification [10]. Monolignols are biosynthesized from phenylalanine (or tyrosine in grasses [11,12]) through a metabolic grid consisting of up to 11 enzyme families and 24 metabolites (Figure 1) [6••,9]. The pathway is initiated by the deamination of phenylalanine or tyrosine and proceeds stepwise through modifications of the phenyl ring and reduction of the propanoid side chains to produce the monolignols (Figure 1). Metabolic flux in monolignol biosynthesis determines lignin properties that affect plant growth, development, and adaptation [2]. Understanding flux regulation is essential for predicting the consequences of genetic perturbations to direct the engineering of specific lignin properties.

With the emergence of high-throughput technologies, it is now possible to describe metabolic flux for lignin biosynthesis by integrating data on gene families, expression patterns of transcripts, proteins, metabolites, and wood phenotypes, at a plant and tissue-specific level [13]. Such an integrative approach applies a systems perspective to biological processes and leads to descriptions of metabolic pathways in predictive models [14]. Only recently has the entire pathway for monolignol biosynthesis been characterized biochemically and genetically in a single tissue of one plant species for modeling [6••]. Metabolic models are fundamental to advanced genetic engineering because of the complexity of the biological systems [9,14].

Metabolic models of monolignol biosynthesis

The major challenge of flux modeling for monolignol biosynthesis is to develop a computational framework to obtain a systematic understanding of the underlying mechanisms regulating the pathway [15•]. Current modeling approaches for metabolic pathways are classified into constraint-based models [16,17], kinetic models [18–20], and black-box models [21].

Figure 1



Flux regulation in monolignol biosynthesis.
Metabolic pathway illustrates general enzyme reactions for monolignol biosynthesis. Red lines represent enzyme inhibitions [36]. Green shading denotes the estimated flux distribution in *P. trichocarpa* [6••,36,43••]. Blue and purple shading represents the G lignin and S lignin-specific metabolic channels, respectively [23,30••]. Abbreviations: phenylalanine ammonia-lyase (PAL); tyrosine ammonia-lyase (TAL); cinnamic acid 4-hydroxylase (C4H); 4-coumaric acid 3-hydroxylase (C3H); caffeoyl shikimate esterase (CSE); hydroxycinnamoyl-CoA shikimate hydroxycinnamoyl transferase (HCT); 4-coumaric acid:CoA ligase (4CL); cinnamoyl-CoA reductase (CCR); cinnamyl alcohol dehydrogenase (CAD); caffeoyl-CoA *O*-methyltransferase (CCoAOMT); coniferaldehyde 5-hydroxylase (CAld5H); 5-hydroxyconiferaldehyde *O*-methyltransferase (AldOMT). Reduced nicotinamide adenine dinucleotide phosphate (NADPH); oxidized nicotinamide adenine dinucleotide phosphate (NADP+); coenzyme A (CoA); adenosine triphosphate (ATP); adenosine monophosphate (AMP); pyrophosphate (PPi); *S*-adenosylmethionine (SAM); *S*-adenosylhomocysteine (SAH).

Constraint-based models

Constraint-based models are ideal when information on kinetic parameters are limited as they only use information about the stoichiometry of the pathway. Models such as Flux Balance Analysis (FBA) assume quasi-steady-state of pathway fluxes that reach equilibrium relatively quickly with respect to varying external fluxes [16,17]. Metabolic pathways are typically overdetermined because the number of reactions exceeds the number of metabolites in the pathways, which means there is no unique solution to these constraint-based problems. Steady-state solutions for constraint-based models are formulated as an optimization problem that relies on a defined set of objective functions (e.g. maximize certain metabolic products such as lignin, or to maintain homeostasis).

Constraint-based models have been developed for monolignol biosynthesis in *Populus* [22], alfalfa [23,24], switchgrass [25], and *Brachypodium* [26]. FBA and a Generalized Mass Action (GMA) approach [27,28] were used to model the monolignol biosynthetic pathway in *Populus* [22].

Kinetic parameters for the GMA model were estimated using the FBA predicted fluxes. The parameters were further optimized to minimize error between predicted and experimental lignin composition for five transgenic *Populus* each downregulated in the expression of a monolignol gene (*AldOMT*, *CAD*, *CAld5H*, *4CL*, and *CCoAOMT*; Figure 1) [22]. The FBA approach was subsequently expanded to explore metabolic regulation and the possibility of metabolic channels in the pathway [24]. Metabolic channels are structures that guide a metabolite product to the next enzyme in a pathway directly, without entering a metabolic pool. Monolignol biosynthesis from young to mature internodes of alfalfa was modeled using FBA coupled with minimization of metabolic adjustment (MOMA) [24] for transgenic lines that independently downregulated the expression of one or more of seven enzymes in the pathway. Variation in lignin composition as stems mature was attributed to flux partitioning at three principal branch points (*p*-coumaroyl-CoA, coniferaldehyde, and coniferyl alcohol) [24]. The model predicted flux control by making assumptions that include the reversibility of enzyme reactions, spatial organization of independent channels for the synthesis of G and S monomers (Figure 1, blue and purple shadings, respectively), and feedback inhibition of S lignin biosynthesis by salicylic acid [24].

To elucidate the presence of metabolic channels, libraries of loosely constrained dynamic models were assembled to represent pathway configurations that include zero, one, or two channels [23]. Model simulations showed that metabolic channels are necessary but insufficient to explain lignin variations in transgenic alfalfa. Crosstalk between S and G pathways may be involved [23]. An FBA model based on transgenic switchgrass independently downregulated in four monolignol enzymes was constructed to test the hypotheses of crosstalk and channeling at diverging points leading to S-monolignol or G-monolignol [25]. Flux distributions based on Monte Carlo simulations supported the possibility of a G lignin-specific channel involving a cinnamoyl-CoA reductase (CCR) and a cinnamyl alcohol dehydrogenase (CAD) (Figure 1, blue shading) [25,29•]. The simulations also identified feedback inhibitions and enzyme competition as essential flux regulators of monolignol biosynthesis in switchgrass [25,29•]. A CCR to CAD channel was proposed for *Brachypodium distachyon* based on steady-state flux analysis using ^{13}C-labeling data [30••]. Heterodimerization of CCR and CAD activates their enzyme activities for monolignol biosynthesis in *Populus trichocarpa* [31].

Despite the aforementioned computational evidence for metabolic channels [23,25,30••], biochemical and enzymatic analyses contradict the assumption of channeling in monolignol biosynthesis. All 24 metabolites could be detected in the lignifying tissues of poplar, suggesting that pathway metabolites are not channeled from one enzyme to another without equilibrating with the bulk solution [32]. Radiotracer assays on reconstituted early pathway enzymes [33] and the C4H–C3H complex [32] showed that substrate channeling does not occur in the conversion of phenylalanine to cinnamic acid, *p*-coumaric acid, and caffeic acid.

Kinetic-based models

Kinetic-based modeling using Michaelis–Menten kinetics is crucial for the quantitative understanding and prediction of flux dynamics in a metabolic pathway containing complex feedback and feedforward regulation [34]. In a kinetic-based approach, reaction fluxes are expressed as a series of rate expressions that are a function of enzyme abundance, substrate concentration, kinetic parameters, and modulators of reactions (e.g. inhibition). The rate expressions are linked through mass conservation to quantify the rate of change of individual metabolite concentrations as a function of the net metabolic flux [27,35]. Although Michaelis–Menten kinetics provide a precise description of reaction flux in a metabolic pathway, the specification of Michaelis–Menten kinetics requires detailed knowledge of the kinetic parameters, which are tedious to obtain [30••]. The monolignol biosynthetic pathway is amenable to kinetic-based modeling because the major components and reactions in the pathway have been identified and can be quantified [6••,36–41].

A Michaelis–Menten kinetic approach was used to model the pathway in stem differentiating xylem (a wood forming tissue) of *P. trichocarpa* based on extensive experimental data [6••,36]. Purified functional recombinant proteins of 21 monolignol enzymes and their isoforms based on genome sequence [42] were assayed for Michaelis–Menten reaction and inhibition kinetics using the 24 pathway metabolites [36]. A comprehensive analysis revealed 104 Michaelis–Menten kinetic parameters and 103 inhibition (Figure 1, red lines) parameters for the stepwise conversion of phenylalanine to monolignols. Protein cleavage-isotope dilution mass spectrometry (PC-IDMS) methods were developed to quantify the absolute protein abundance of each of the multiple isoforms of the pathway enzymes in stem differentiating xylem [36,38]. To predict the cooperative regulation of flux dynamics and distribution (Figure 1, green shading) leading to lignin content and composition, the 207 kinetic parameters and the abundance of proteins were formulated into a kinetic model containing 24 mass balance equations [6••,36]. The kinetic model revealed novel features and mechanisms of regulation. For example, most enzymes in the pathway are produced in excess of what is required for a normal lignin phenotype [36]. The model explained a long-standing paradox regarding the regulation of monolignol subunit ratios in lignin, revealing that pathway flux is preferentially diverged into S-lignin, and that early-pathway enzymes control S/G

ratio by modulating the metabolic flux for G-lignin. The model also showed that metabolic channels are not necessary for regulating the S/G ratio in *P. trichocarpa* [36].

Kinetic modeling typically assumes enzymes to be functionally independent. However, for monolignol biosynthesis, protein–protein interactions have already been identified that impart significant regulatory control over enzyme activity and metabolic flux in the pathway [32,37,43••,44]. Two multimeric protein complexes have been characterized in *P. trichocarpa*. A complex of two isoforms of cinnamic acid 4-hydroxylases (C4H) and a *p*-coumarate ester 3-hydroxylase (C3H) exhibit enzymatic efficiencies (V_{max}/K_m) that are 70–6500 times greater than the efficiency of individual enzymes [32]. A tetrameric complex containing two isoforms of 4-coumaric acid:CoA ligases (4CL3 and 4CL5) regulates the specific inhibition and activation of metabolic fluxes that convert hydroxycinnamic acids to their CoA thioester derivatives [37].

A kinetics approach was used to model the regulatory function of the 4CL3–4CL5 complex [37,43••]. The model was derived from experimentally verified mechanistic interactions and kinetic parameters between the enzymes and substrates that participate in the CoA ligations of *p*-coumaric acid and caffeic acid. The interaction of 4CL3 and 4CL5 has a functional role with 4CL5 acting to control the total reaction rates of CoA-ligation [37]. The respective competitive, uncompetitive, and self-inhibitions of 4CL3 and 4CL5 are necessary inclusions in the model to more closely fit the mixed enzyme–mixed substrate reaction assays [37]. The 4CL3–4CL5 complex also produces significant variation in the steady-state flux of the pathway. The enzyme complex increased the stability of the pathway from 70% to 92% by activating redundant pathways for the CoA ligation step in the biosynthesis of S and G monolignols [43••]. The system switches to these alternate pathways under extreme perturbations. In the presence of the complex, metabolic flux for monolignol biosynthesis is sensitive to changes in the ratio of 4CL3 and 4CL5 and can be effectively modulated by altering the abundance of 4CL5 [43••].

Black-box models

In addition to Michaelis–Menten kinetics, methods based on machine-learning optimization [21] may allow the deduction of parameters for modeling metabolic pathways. An advantage of machine learning approaches is that there are no underlying assumptions about the data. The models can handle linear and non-linear data as well as different types of data (discrete, continuous, and factors). In machine learning approaches, several hyperparameters need to be tuned to maximize performance. These models tend to overfit the data and require post-processing to reduce the variance in model predictions. In addition, these models are a black box with limited ability to infer biological significance. A machine learning model based on Artificial Neural Networks (ANN) [45] was assembled (Punith Naik, PhD thesis, NC State University, 2016) to predict lignin structure using proteomics and transcriptomics data from transgenic and wildtype *P. trichocarpa* [6••]. The hyperparameters of the neural network, such as the number of nodes and number of hidden layers, were estimated using cross-validation. The network architecture that results in the minimum cross-validation error is the optimal solution. The ANN-based model (Punith Naik, PhD thesis, NC State University, 2016) was able to account for 74% of the variation in lignin S/G ratios across 221 independent transgenic lines.

Applications of modeling for lignin engineering

Many transgenic and mutant plants with modified expression of monolignol genes have been produced to understand the genetic regulation of lignin, but results are generally unpredictable [7–10,46,47,48•,49]. To gain a predictive understanding of how changing gene expression affects lignin and wood properties, an integrative systems model was recently assembled using multi-omics data from ~2000 transgenic *P. trichocarpa* systematically perturbed in the expression of 21 monolignol pathway genes [6••]. The model integrated five levels of omics data: genomics, transcriptomics, proteomics, metabolomics (fluxomics), and phenomics (25 traits including lignin content and composition, stem-volume, density, mechanical strength, and saccharification efficiency) to describe the transduction of regulatory information from genotype to phenotype for lignin biosynthesis [6••]. The model predicts how changing the expression of any pathway gene or gene combination affects these 25 traits, enabling the directed engineering of these traits individually or in combinations. Simulations of every possible combination of monolignol gene perturbations identified optimal strategies that could simultaneously improve multiple lignin and wood properties (e.g. reduce lignin, increase density, and elevate saccharification efficiency) while minimizing negative effects on growth [6••]. These multi-gene engineering strategies, however, remain to be confirmed *in vivo* and under diverse environmental conditions to improve the model.

Outlook

Increasingly accurate models that can predict how genetic modifications alter the physical and chemical properties of lignin must incorporate factors beyond enzymatic parameters. Additional factors include incorporating spatial and stochastic modeling [30••,50] to the current models and connecting the metabolic models with mathematical descriptions at different layers of biological regulation. These regulations may include transcriptional networks that control monolignol gene expressions [51], post-translational modification that modulates pathway enzyme activity [52], and lignin polymerization [53]. Accounting for enzymes such as glucosyltransferases [54], glucosidases [55], and hydroxycinnamaldehyde dehydrogenases [56],

which metabolize monolignol precursors, may further increase model predictive power. The spatial [57,58] and temporal [59] expression of monolignol genes could be modeled to understand lignification in different cell-types, tissues, and developmental stages. Cell and tissue models of lignin biosynthesis will be important tools to strategically develop designer plants [57,58]. There is, for example, evidence of spatial differences in the locations of the alcohol and aldehyde groups of the polymerized lignin in the secondary cell wall of fiber cells of bristle grass, but not in *Arabidopsis thaliana* [60]. Incorporating regulation at the transcriptional and post-translational levels and modeling the polymerization process are important next steps to advance predictive understanding of the pathway for lignin engineering.

Conflict of interest statement

Nothing declared.

Acknowledgements

This article was supported by the National Natural Science Foundation of China (NSFC) Grants 31430093 and 31470672. We also thank the financial support from the U.S. National Science Foundation, Plant Genome Research Program Grant DBI-0922391, the NC State University Jordan Family Distinguished Professor Endowment, and the NC State University Forest Biotechnology Industrial Research Consortium, and the 111 Project (B16010) of the Ministry of Education of China.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Bar-On YM, Phillips R, Milo R: **The biomass distribution on Earth**. *Proc Natl Acad Sci U S A* 2018, **115**:6506-6511 http://dx.doi.org/10.1073/pnas.1711842115.

2. Sarkanen KV, Ludwig CH: *Lignins: Occurrence, Formation, Structure and Reactions*. Wiley-Interscience; 1971.

3. Freudenberg K: **Lignin: its constitution and formation from *p*-hydroxycinnamyl alcohols: lignin is duplicated by dehydrogenation of these alcohols; intermediates explain formation and structure**. *Science* 1965, **148**:595-600.

4. Martinez AT, Rencoret J, Marques G, Gutierrez A, Ibarra D, Jimenez-Barbero J, del Rio JC: **Monolignol acylation and lignin structure in some nonwoody plants: a 2D NMR study**. *Phytochemistry* 2008, **69**:2831-2843.

5. Voelker SL, Lachenbruch B, Meinzer FC, Kitin P, Strauss SH: **Transgenic poplars with reduced lignin show impaired xylem conductivity, growth efficiency and survival**. *Plant Cell Environ* 2011, **34**:655-668.

6. •• Wang JP, Matthews ML, Williams CM, Shi R, Yang C, Tunlaya-Anukit S, Chen H-C, Li Q, Liu J, Lin C-Y *et al.*: **Improving wood properties for wood utilization through multi-omics integration in lignin biosynthesis**. *Nat Commun* 2018, **9**:1579.
The authors present a comprehensive systems model of lignin biosynthesis based on multi-omics data from ~2000 transgenic poplars. The model predicts the outcomes of modifying multiple genes involved in monolignol biosynthesis on 25 wood traits, and predicts ways to strategically engineer wood for improved properties and utilization.

7. Boerjan W, Ralph J, Baucher M: **Lignin biosynthesis**. *Annu Rev Plant Biol* 2003, **54**:519-546.

8. Dixon RA, Reddy MSS, Gallego-Giraldo L: **Monolignol biosynthesis and its genetic manipulation: the good, the bad, and the ugly**. *Rec Adv Polyphen Res* 2014, **vol. 4**1-38.

9. Umezawa T: **Lignin modification in planta for valorization**. *Phytochem Rev* 2018, **17**:1305-1327.

10. Sederoff RR, MacKay JJ, Ralph J, Hatfield RD: **Unexpected variation in lignin**. *Curr Opin Plant Biol* 1999, **2**:145-152.

11. Maeda HA: **Lignin biosynthesis: tyrosine shortcut in grasses**. *Nat Plants* 2016, **2**:16080.

12. Neish AC: **Formation of meta-coumaric and para-coumaric acids by enzymatic deamination of the corresponding isomers of tyrosine**. *Phytochemistry* 1961, **1**:1-24.

13. Salon C, Avice J-C, Colombié S, Dieuaide-Noubhani M, Gallardo K, Jeudy C, Ourry A, Prudent M, Voisin A-S, Rolin D: **Fluxomics links cellular functional analyses to whole-plant phenotyping**. *J Exp Bot* 2017, **68**:2083-2098.

14. Sweetlove LJ, Nielsen J, Fernie AR: **Engineering central metabolism - a grand challenge for plant biologists**. *Plant J* 2017, **90**:749-763.

15. • Voit EO: **The best models of metabolism**. *Wiley Interdiscip Rev: Syst Biol Med* 2017, **9**.
A comprehensive review of mathematical modeling approaches for metabolic pathways covering the design process, modeling frameworks, and computational methods for metabolic models.

16. Schilling CH, Palsson BO: **Assessment of the metabolic capabilities of *Haemophilus influenzae* Rd through a genome-scale pathway analysis**. *J Theor Biol* 2000, **203**:249-283.

17. Schuster S, Dandekar T, Fell DA: **Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering**. *Trends Biotechnol* 1999, **17**:53-60.

18. Grimbs S, Selbig J, Bulik S, Holzhutter HG, Steuer R: **The stability and robustness of metabolic states: identifying stabilizing sites in metabolic networks**. *Mol Syst Biol* 2007, **3**:146.

19. Palsson BO, Lightfoot EN: **Mathematical modelling of dynamics and control in metabolic networks. I. On Michaelis-Menten kinetics**. *J Theor Biol* 1984, **111**:273-302.

20. Reich JG, Sel'kov EE: *Energy Metabolism of the Cell: A Theoretical Treatise*. Academic Press; 1981.

21. Sriyudthsak K, Shiraishi F, Hirai MY: **Mathematical modeling and dynamic simulation of metabolic reaction systems using metabolome time series data**. *Front Mol Biosci* 2016, **3**:15.

22. Lee Y, Voit EO: **Mathematical modeling of monolignol biosynthesis in *Populus* xylem**. *Math Biosci* 2010, **228**:78-89.

23. Lee Y, Escamilla-Trevino L, Dixon RA, Voit EO: **Functional analysis of metabolic channeling and regulation in lignin biosynthesis: a computational approach**. *PLoS Comput Biol* 2012, **8**:e1002769.

24. Lee Y, Chen F, Gallego-Giraldo L, Dixon RA, Voit EO: **Integrative analysis of transgenic alfalfa (*Medicago sativa* L.) suggests new metabolic control mechanisms for monolignol biosynthesis**. *PLoS Comput Biol* 2011, **7**:e1002047.

25. Faraji M, Fonseca LL, Escamilla-Trevino L, Dixon RA, Voit EO: **Computational inference of the structure and regulation of the lignin pathway in *Panicum virgatum***. *Biotechnol Biofuels* 2015, **8**:151.

26. Faraji M, Fonseca LL, Escamilla-Trevino L, Barros-Rios J, Engle NL, Yang ZK, Tschaplinski TJ, Dixon RA, Voit EO: **A dynamic model of lignin biosynthesis in Brachypodium distachyon**. *Biotechnol Biofuels* 2018, **11**:253.

27. Horn F, Jackson R: **General mass action kinetics**. *Arch Ration Mech Anal* 1972, **47**:81-116.

28. Voit EO: **Biochemical systems theory: a review**. *J ISRN Biomath* 2013, **2013** 897658.

29. • Faraji M, Voit EO: **Improving bioenergy crops through dynamic metabolic modeling**. *Processes* 2017, **5**.
A review of computational techniques for dynamic modeling of plant metabolism. A case study for switchgrass lignin biosynthesis was presented to demonstrate the potential for model-guided perturbation of lignin content and composition.

30. •• Faraji M, Fonseca LL, Escamilla-Trevino L, Barros-Rios J, Engle N, Yang ZK, Tschaplinski TJ, Dixon RA, Voit EO: **Mathematical models of lignin biosynthesis**. *Biotechnol Biofuels* 2018, **11**.
Comparative analysis of computational models for lignin biosynthesis in poplar, alfalfa, switchgrass, and*Brachypodium distachyon*. The authors show that different plant species have distinct regulatory features and spatiotemporal characteristics in lignin biosynthesis, requiring species-specific models to describe the pathway.

31. Yan X, Liu J, Kim H, Liu B, Huang X, Yang Z, Lin YC, Chen H, Yang C, Wang JP *et al.*: **CAD1 and CCR2 protein complex formation in monolignol biosynthesis in Populus trichocarpa**. *New Phytol* 2018 http://dx.doi.org/10.1111/nph.15505.

32. Chen HC, Li QZ, Shuford CM, Liu J, Muddiman DC, Sederoff RR, Chiang VL: **Membrane protein complexes catalyze both 4-and 3-hydroxylation of cinnamic acid derivatives in monolignol biosynthesis**. *Proc Natl Acad Sci U S A* 2011, **108**:21253-21258.

33. Ro DK, Douglas CJ: **Reconstitution of the entry point of plant phenylpropanoid metabolism in yeast (Saccharomyces cerevisiae): implications for control of metabolic flux into the phenylpropanoid pathway**. *J Biol Chem* 2004, **279**:2600-2607.

34. Schallau K, Junker BH: **Simulating plant metabolic pathways with enzyme-kinetic models**. *Plant Physiol* 2010, **152**:1763-1771.

35. Savageau MA: *Biochemical Systems Analysis: A Study of Function and Design in Molecular Biology*. Reading, Mass: Addison-Wesley Pub. Co., Advanced Book Program; 1976.

36. Wang JP, Naik PP, Chen HC, Shi R, Lin CY, Liu J, Shuford CM, Li QZ, Sun YH, Tunlaya-Anukit S *et al.*: **Complete proteomic-based enzyme reaction and inhibition kinetics reveal how monolignol biosynthetic enzyme families affect metabolic flux and lignin in *Populus trichocarpa***. *Plant Cell* 2014, **26**:894-914.

37. Chen HC, Song J, Wang JP, Lin YC, Ducoste J, Shuford CM, Liu J, Li QZ, Shi R, Nepomuceno A *et al.*: **Systems biology of lignin biosynthesis in *Populus trichocarpa*: heteromeric 4-coumaric acid: coenzyme A ligase protein complex formation, regulation, and numerical modeling**. *Plant Cell* 2014, **26**:876-893.

38. Shuford CM, Li Q, Sun YH, Chen HC, Wang J, Shi R, Sederoff RR, Chiang VL, Muddiman DC: **Comprehensive quantification of monolignol-pathway enzymes in *Populus trichocarpa* by protein cleavage isotope dilution mass spectrometry**. *J Proteome Res* 2012, **11**:3390-3404.

39. Osakabe K, Tsao CC, Li L, Popko JL, Umezawa T, Carraway DT, Smeltzer RH, Joshi CP, Chiang VL: **Coniferyl aldehyde 5-hydroxylation and methylation direct syringyl lignin biosynthesis in angiosperms**. *Proc Natl Acad Sci U S A* 1999, **96**:8955-8960.

40. Li L, Popko JL, Umezawa T, Chiang VL: **5-Hydroxyconiferyl aldehyde modulates enzymatic methylation for syringyl monolignol formation, a new view of monolignol biosynthesis in angiosperms**. *J Biol Chem* 2000, **275**:6537-6545.

41. Lin CY, Wang JP, Li Q, Chen HC, Liu J, Loziuk P, Song J, Williams C, Muddiman DC, Sederoff RR *et al.*: **4-Coumaroyl and caffeoyl shikimic acids inhibit 4-coumaric acid:coenzyme A ligases and modulate metabolic flux for 3-hydroxylation in monolignol biosynthesis of *Populus trichocarpa***. *Mol Plant* 2015, **8**:176-187.

42. Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A *et al.*: **The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray)**. *Science* 2006, **313**:1596-1604.

43. •• Naik P, Wang JP, Sederoff R, Chiang V, Williams C, Ducoste JJ: **Assessing the impact of the 4CL enzyme complex on the robustness of monolignol biosynthesis using metabolic pathway analysis**. *PLoS One* 2018, **13**:e0193896.
Computational models and Monte Carlo simulations were used to investigate the role of Ptr4CL3 and Ptr4CL5 enzyme complex formation on the steady state flux distribution in the monolignol pathway. Robustness and stability analyses showed that the enzyme complex increases the stability of pathway flux by 22%.

44. Gou M, Ran X, Martin DW, Liu CJ: **The scaffold proteins of lignin biosynthetic cytochrome P450 enzymes**. *Nat Plants* 2018, **4**:299-310.

45. Schalkoff RJ: *Artificial Neural Networks*. McGraw-Hill; 1997.

46. Elkind Y, Edwards R, Mavandad M, Hedrick SA, Ribak O, Dixon RA, Lamb CJ: **Abnormal plant development and down-regulation of phenylpropanoid biosynthesis in transgenic tobacco containing a heterologous phenylalanine ammonia-lyase gene**. *Proc Natl Acad Sci U S A* 1990, **87**:9057-9061.

47. Hu WJ, Harding SA, Lung J, Popko JL, Ralph J, Stokke DD, Tsai CJ, Chiang VL: **Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees**. *Nat Biotechnol* 1999, **17**:808-812.

48. • Wang Jack P, Tunlaya-Anukit S, Shi R, Yeh TF, Chuang L, Isik F, Yang C, Liu J, Li Q, Loziuk Philip L *et al.*: **A proteomic-based quantitative analysis of the relationship between monolignol biosynthetic protein abundance and lignin content using transgenic *Populus trichocarpa***. Edited by Quideau S, Yoshida K. 2016, **vol. 5**:89-107.
Mathematical models were constructed to analyze the quantitative relationship between monolignol protein abundance and lignin content using 80 transgenic poplar lines individually downregulated in the expression of 5 monolignol pathway enzyme families. The models suggest that monolignol proteins are present in excess to a quantity needed for normal lignin content.

49. Li L, Zhou YH, Cheng XF, Sun JY, Marita JM, Ralph J, Chiang VL: **Combinatorial modification of multiple lignin traits in trees through multigene cotransformation**. *Proc Natl Acad Sci U S A* 2003, **100**:4939-4944.

50. Islam MM, Saha R: **Computational approaches on stoichiometric and kinetic modeling for efficient strain design**. *Methods Mol Biol* 2018, **1671**:63-82.

51. Lin YJ, Chen H, Li Q, Li W, Wang JP, Shi R, Tunlaya-Anukit S, Shuai P, Wang Z, Ma H *et al.*: **Reciprocal cross-regulation of VND and SND multigene TF families for wood formation in *Populus trichocarpa***. *Proc Natl Acad Sci U S A* 2017, **114**:E9722-E9729.

52. Wang JP, Chuang L, Loziuk PL, Chen H, Lin YC, Shi R, Qu GZ, Muddiman DC, Sederoff RR, Chiang VL: **Phosphorylation is an on/off switch for 5-hydroxyconiferaldehyde *O*-methyltransferase activity in poplar monolignol biosynthesis**. *Proc Natl Acad Sci U S A* 2015, **112**:8481-8486.

53. Lin CY, Li QZ, Tunlaya-Anukit S, Shi R, Sun YH, Wang JP, Liu J, Loziuk P, Edmunds CW, Miller ZD *et al.*: **A cell wall-bound anionic peroxidase, PtrPO21, is involved in lignin polymerization in *Populus trichocarpa***. *Tree Genet Genomes* 2016, **12**.

54. Lim EK, Li Y, Parr A, Jackson R, Ashford DA, Bowles DJ: **Identification of glucosyltransferase genes involved in sinapate metabolism and lignin synthesis in Arabidopsis**. *J Biol Chem* 2001, **276**:4344-4349.

55. Chapelle A, Morreel K, Vanholme R, Le-Bris P, Morin H, Lapierre C, Boerjan W, Jouanin L, Demont-Caulet N: **Impact of the absence of stem-specific beta-glucosidases on lignin and monolignols**. *Plant Physiol* 2012, **160**:1204-1217.

56. Nair RB, Bastress KL, Ruegger MO, Denault JW, Chapple C: **The Arabidopsis thaliana REDUCED EPIDERMAL FLUORESCENCE1 gene encodes an aldehyde dehydrogenase involved in ferulic acid and sinapic acid biosynthesis**. *Plant Cell* 2004, **16**:544-554.

57. Shi R, Wang JP, Lin YC, Li QZ, Sun YH, Chen H, Sederoff RR, Chiang VL: **Tissue and cell-type co-expression networks of transcription factors and wood component genes in *Populus trichocarpa***. *Planta* 2017, **245**:927-938.

58. Le Roy J, Blervacq AS, Creach A, Huss B, Hawkins S, Neutelings G: **Spatial regulation of monolignol biosynthesis and laccase genes control developmental and stress-related lignin in flax**. *BMC Plant Biol* 2017, **17**:124.

59. Liu J, Hai G, Wang C, Cao S, Xu W, Jia Z, Yang C, Wang JP, Dai S, Cheng Y: **Comparative proteomic analysis of *Populus trichocarpa* early stem from primary to secondary growth**. *J Proteomics* 2015, **126**:94-108.

60. Liu B, Wang P, Kim JI, Zhang D, Xia Y, Chapple C, Cheng JX: **Vibrational fingerprint mapping reveals spatial distribution of functional groups of lignin in plant cell wall**. *Anal Chem* 2015, **87**:9436-9442.