in silico
Plants
OXFORD

Original Research

AraRoot—a comprehensive genome-scale metabolic model for the Arabidopsis root system

Lohani Esterhuizen[1], Nicholas Ampimah[1], Marna D. Yandeau-Nelson[2], Basil J. Nikolau[3],
Erin E. Sparks[4], , Rajib Saha[1,*],

[1]Chemical and Biomolecular Engineering, University of Nebraska-Lincoln, Lincoln, NE, United States
[2]Department of Genetics, Development and Cell Biology, Iowa State University, Ames, IA, United States
[3]Roy J. Carver Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University, Ames, IA, United States
[4]Department of Plant and Soil Sciences, University of Delaware, Newark, DE, United States

*Corresponding author. Chemical and Biomolecular Engineering, University of Nebraska-Lincoln, Lincoln, NE, United States. E-mail: rsaha2@unl.edu

Handling Editor: Amy Marshall-Colon

Abstract Being the first plant to have its genome sequenced, *Arabidopsis thaliana* (Arabidopsis) is a well-established genetic model plant system. Studies on Arabidopsis have provided major insights into the physiological and biochemical nature of plants. Methods that allow us to study organisms' metabolism computationally include using genome-scale metabolic models (GEMs). Despite its popularity, no GEM currently maps the metabolic activity in the roots of Arabidopsis, which is the organ that faces and responds to stress conditions in the soil. We have developed a comprehensive metabolic model of the Arabidopsis root system—AraRoot. The final model includes 2,682 reactions, 2,748 metabolites, and 1,310 genes. Analyzing the metabolic pathways in this model identified 158 possible bottleneck genes that impact biomass production, most of which were found to be related to phosphorous-containing- and energy-related pathways. Further insights into tissue-specific metabolic reprogramming conclude that the cortex layer in the roots is likely responsible for root growth under prolonged exposure to high salt conditions. At the same time, the endodermis and epidermis are responsible for producing metabolites responsible for increased cell wall biosynthesis. The epidermis was found to have a very poor ability to regulate its metabolism during exposure to high salt concentrations. Overall, AraRoot is the first metabolic model that comprehensively captures the biomass formation and stress responses of the tissues in the Arabidopsis root system.

KEYWORDS: AraRoot; Arabidopsis; root system; tissue-specific; genome-scale metabolic model.

1 INTRODUCTION

It is well-established that *Arabidopsis thaliana* (Arabidopsis) is a popular model organism in plant biology (The Arabidopsis Genome Initiative, 2000; Chen *et al.*, 2004; Brady *et al.*, 2007; W. Xu *et al.*, 2013; Smolko *et al.*, 2021). Although not economically important, its simple metabolism, short lifespan, and minimal nutrient requirements have made it the ideal model organism to study plant molecular, cellular, and developmental mechanisms (The Arabidopsis Genome Initiative, 2000; Poolman *et al.*, 2009). The sequencing of the Arabidopsis genome provided much-needed insights into plant metabolism, with recent studies focusing on tissue-specific metabolic pathways (Mintz-Oron *et al.*, 2012; de Oliveira Dal'Molin *et al.*, 2015a; Shaw & Cheung, 2018). When exposing Arabidopsis to altered growth media and stress conditions, the root system detects and responds to changes in the soil, such as water scarcity and nutrient depletion (Kellermeier *et al.*, 2014; Smolko *et al.*, 2021; H. Xu *et al.*,

2022; Ngo & Nakamura, 2022). The main functionality of the roots is to take up all required nutrients and water from the soil while also acting as an anchor and providing some storage for excess nutrients (Smolko *et al.*, 2021; Chowdhury *et al.*, 2022). The absence of complex metabolic processes such as photosynthesis in the roots coupled with a simple cellular organization make this an excellent organ for metabolic modeling (Brady *et al.*, 2007; W. Xu *et al.*, 2013; Kellermeier *et al.*, 2014).

As novel plant biology discoveries have continued to be made, scientists looked to the power of machines for help, which brought about the rise of the genome-scale metabolic model (GEM) (Varma & Palsson, 1994; C. G. de O. Dal'Molin et al., 2010; Thiele & Palsson, 2010; Saha et al., 2011; C. G. O. Dal'Molin et al., 2014; de Oliveira Dal'Molin *et al.*, 2015; Metcalf et al., 2020). A GEM mathematically formulates the metabolism of an organism where a matrix represents the stoichiometry between the metabolites and the reactions involved in known biological processes as mapped through genome studies, with

associated genes mapped through gene-protein-reaction (GPR) relations (Edwards & Palsson, 1999; Mahadevan et al., 2002; Metcalf et al., 2020). A GEM captures a vast number of metabolic pathways. In any GEM, a metabolite typically participates in more than one reaction. With the reaction fluxes in a GEM being the unknowns, this high dimensionality and metabolic redundancy would make a GEM an underdetermined system, which would have an infinite number of feasible solutions (Edwards & Palsson, 1999; Poolman et al., 2009). A method to reduce the solution space of the model is to introduce 'omics' data to the model to restrict the bounds of the reaction fluxes in the model (Gudmundsson & Thiele, 2010; Sinha et al., 2021).

Experimentally obtained omics data such as transcriptomics and proteomics data can be incorporated into the model through the GPR by using any one of the well-established methods. Some well-known examples include GIMME (Becker & Palsson, 2008), iMAT (Zur et al., 2010), and E-Flux (Colijn et al., 2009). The expression distributed reaction flux measurement (EXTREAM) algorithm is built on the E-Flux method to account for gene abundance in the organism (Chowdhury et al., 2023). A study conducted by Machado and Herrgård (2014) compared 18 different published methods that incorporate transcriptomics into several different GEMs and concluded that in the case of a biomass objective function, E-Flux is the only method that does not replace the biomass objective with an objective that relies on the specific set of gene expression data (Machado & Herrgård, 2014) and EXTREAM only has a significant impact on metabolic predictions in the case of complex gene relations.

Though experimentally determined biomass composition information yields the most accurate growth predictions by a GEM, it is expensive and time-consuming to conduct these experiments, which is why GEMs often assume a biomass composition established for other similar organisms (de Oliveira Dal'Molin *et al.* 2015; Chowdhury et al., 2022). The maize root model published by Chowdhury et al. (2022) included a biomass reaction that captured the stoichiometry of each metabolite that contributed to biomass formation based on experimental evidence obtained specifically for the maize root system (Chowdhury et al., 2022). The tissue-specific model published by Schroeder and Saha (2020a) contained an Arabidopsis root model consisting of a basic biomass composition that included only five metabolites, based on the composition of switchgrass (*Panicum virgatum*) (Schroeder and Saha 2020a, b).

Although tissue-specific models of Arabidopsis have been developed for whole-plant metabolism, there is currently no tissue-specific metabolic model for the Arabidopsis root system that comprehensively captures biomass composition and specialized metabolism. This work aims to develop a comprehensive metabolic model for the Arabidopsis root system—AraRoot. Figure 1 summarizes the methodology followed to reconstruct AraRoot, the first metabolic model capable of capturing the biomass formation and stress response of the Arabidopsis root system as determined from previous literature (Dinneny *et al.*, 2008; Iyer-Pascuzzi *et al.*, 2011; Geng *et al.*, 2013; Han & Yang, 2021; Smolko *et al.*, 2021). A basic local alignment search tool for proteins (BLASTp) (Altschul *et al.*, 1990) confirmed high gene homology between Arabidopsis and maize (Saha *et al.*, 2011; Provart *et al.*, 2016; Woodhouse *et al.*, 2021), motivating the use

of a published maize root model from Chowdhury *et al.* as a blueprint. In addition, AraRoot includes data from established databases such as KEGG (Kanehisa *et al.*, 2014), TAIR (Berardini *et al.*, 2015), and MaizeGDB (Woodhouse *et al.*, 2021). The model was updated to include a comprehensive biomass reaction and specialized metabolites such as dipeptides, coumarins, and lignols, which include coniferin, esculetin, scopoletin, and scopoline as described by literature (Strehmel *et al.*, 2014). Results indicate that sugar components act as overflow metabolites for biomass formation, while cellulose and xyloglucan concentrations have the highest impact on biomass synthesis (Zabotina *et al.*, 2012). A new in-house tool known as OptRecon was used to remove thermodynamically infeasible cycles (TICs) and to add missing pertinent reactions from a custom database while preventing the formation of new TICs. Root tissue transcriptomic data was integrated using the E-Flux algorithm (Li *et al.*, 2016) to reduce the solution space and contextualize the root model. Our recently developed metabolic bottleneck analysis (MBA) was next applied to the model to identify the bottleneck reactions (and the associated bottleneck genes) that have an impact on biomass production when strict bounds are imposed on the reaction fluxes (Chowdhury *et al.*, 2023).

Published gene expression data (Geng *et al.*, 2013) was utilized to create cell-type-specific models of the cortex, endodermis, epidermis, and stele tissues under normal and stress conditions using the AraRoot model as a blueprint. Subsequently, the same transcriptomics data were used to analyze the bottleneck genes, thus identifying correlations between the expression of bottleneck genes under stress conditions to limits in biomass production. A flux range variation analysis of each cell type under normal and high salt growth conditions captured the metabolic reprogramming of each tissue during stress-induced conditions. Resulting variations in reaction flux ranges indicate that the cortex ultimately drives growth recovery in the roots after prolonged exposure to salt, with increased reaction flux ranges found in most of the metabolic pathways (Geng *et al.*, 2013). The epidermis and endodermis were found to be responsible for producing more metabolites related to cell wall biosynthesis with flux ranges in the pentose phosphate pathway increasing during salt exposure, particularly the production of D-ribulose 5-phosphate (Huang *et al.*, 2020; H. Xu *et al.*, 2022). The epidermis was found to lack the ability to further regulate metabolism during stress-induced growth (Dinneny *et al.*, 2008; Geng *et al.*, 2013). Together, this work proposes a tissue-specific metabolic model of the roots, capable of capturing metabolic variation under different growth conditions.

2. METHODS

2.1 Initializing the model

The reconstruction and analysis of the AraRoot model aim to produce a metabolic model capable of comprehensively capturing the biomass formation and stress response of the Arabidopsis root system as described in the literature (Dinneny *et al.*, 2008; Iyer-Pascuzzi *et al.*, 2011; Geng *et al.*, 2013; Smolko *et al.*, 2021). Model initialization involves establishing a base model capable of reconstruction to represent an Arabidopsis root system. Previous literature indicates high genetic similarity between
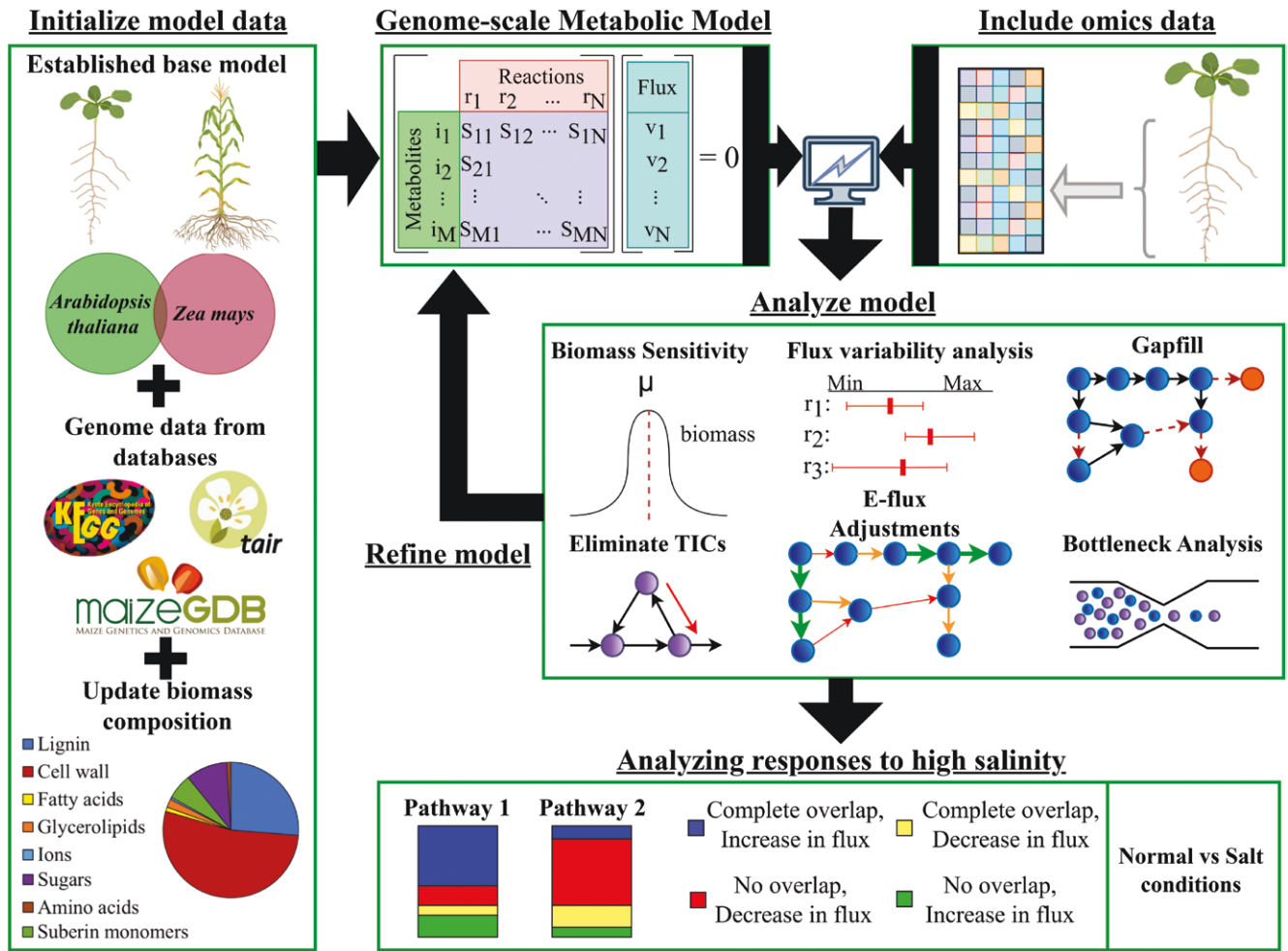
Figure 1. Reconstructing and analyzing the AraRoot model by (A) updating a base root model (*Zea mays*) to include comprehensive Arabidopsis root biomass data and reactions from well-known databases, (B) constructing the base model and including published transcriptomics into the model, which was analyzed using established methods, and (C) analyzing the response of the model to stress-induced environmental conditions by comparing the flux range variations under normal and salt stress conditions available from published data (Kiegle *et al.*, 2000; Geng *et al.*, 2013; Han & Yang, 2021; Smolko *et al.*, 2021).

maize and Arabidopsis with an estimated 90% of the proteins found in maize having homologs in Arabidopsis (Brendel *et al.*, 2002; Saha *et al.*, 2011; Chowdhury *et al.*, 2022). The protein sequences of genes in the Arabidopsis root system, as established by Li *et al.* (2016), were obtained from the TAIR database and were set as the query sequences for BLASTp. The protein sequences of the genes included in the published maize root base model were obtained from the KEGG database and were considered the subject sequences. The screening criteria for the BLASTp result hits were set at a percentage identity (PID) of 60%, a query coverage (COV) of 0.8, and a Max Score of 400.

The biomass composition of the Arabidopsis root system was determined by consulting literature. A complete list of all metabolites and their respective compositions that comprise the biomass can be found in Supplementary File 1, accompanied by links to the literature sources. The biomass reaction of the root system in the base maize model was updated according to the new Arabidopsis biomass to accurately represent the Arabidopsis root system. Several well-known techniques were used to ensure that the base model could produce biomass with

the updated composition. These techniques include flux balance analysis (FBA) and flux variability analysis (FVA). The mathematical formulation of the optimization problem used in both techniques can be found in Supplementary File 2.

A previously published algorithm was used to analyze the sensitivity of the model to perturbations in the updated biomass metabolite coefficients (Dinh *et al.*, 2022). The proposed algorithm calculates a standard deviation ratio (SDR), which quantitively determines the impact that deviations in the input parameter c have on the output fluxes v of a stoichiometric metabolic model using the following equation:

$$SDR_c^v = \frac{\Delta(\sigma_v \backslash \mu_v)}{\Delta(\sigma_c \backslash \mu_c)} \tag{1}$$

Where σ and μ represents the standard deviation and mean, respectively. In Equation 1 $\Delta(\sigma_v \backslash \mu_v)$ represents the change in the relative standard deviation of the output (biomass yield) that results from introducing noise to the system, and similarly $\Delta(\sigma_c \backslash \mu_c)$ represent the change in the relative standard deviation

of the input parameter (biomass coefficient) as subject to the introduction of noise (MacGillivray *et al.*, 2017). Parameters were sampled 10,000 times from a normal distribution with relative standard deviations of 5%, 10%, and 20% centered around the original values. Noise was imposed on all coefficients simultaneously, followed by imposition on cellulose and xylan singularly as biomass metabolites.

For further insights into the relationship between the metabolites and biomass formation in the base model, the shadow price values (λ) of the metabolites were calculated by considering the dual problem of the FBA formulation. The shadow price value calculated for each metabolite quantifies the marginal sensitivity of biomass (objective function) to single-unit perturbations in the availability of a limiting metabolite (Varma & Palsson, 1994). An upper and lower shadow price range was determined by utilizing the coefficients of the biomass reactions determined at 20% perturbations during the biomass sensitivity analysis. Negative shadow price values would indicate growth-limiting metabolites, which are metabolites whose availability indirectly affects biomass formation. Positive shadow price values are associated with overflow metabolites, which are metabolites whose availability directly affects biomass formation. The mathematical formulation for calculating the shadow price value of each metabolite can be found in Supplementary File 2.

2.2 Model reconstruction

After incorporating the updated biomass reaction into the base model, a standard model reconstruction procedure was followed with added steps to capture comprehensive root metabolism (Thiele & Palsson, 2010). It was necessary to determine which reactions in the base model had become redundant given the new biomass reaction since some reactions in the model were only specific to maize and would also need to be removed. FVA was utilized to identify all blocked reactions, which are reactions that did not carry flux during biomass production. The blocked reactions were compared to reactions catalyzed by specific gene products that showed high sequence homology between Arabidopsis and maize. Common reactions remained in the model, while blocked reactions with no homology between maize and Arabidopsis were removed from the model.

An in-house tool known as OptRecon was utilized to identify and remove TICs (Nelson & Saha, 2024). The mathematical formulation capable of identifying TICs in a GEM is provided in Supplemental File 2 and forms the basis for OptRecon. In addition to identifying TICs, OptRecon is also able to add reactions to any GEM in such a way as to not create any new TICs. To account for any reactions that were missing from the maize base model but were relevant to the Arabidopsis root system, a custom database of reactions was created from which relevant reactions were added to AraRoot. The custom database was created by comparing the transcriptomes of the Arabidopsis root system (Li *et al.*, 2016) to the published root system transcriptomes of other similar C3 plants, including tomato (Pirona *et al.*, 2023), soybean (Adhikari *et al.*, 2019), and rice (Liu *et al.*, 2021). Reactions that were found to have high homology between Arabidopsis and the three other plants were compiled to form a custom database, and OptRecon was utilized to incorporate the maximum number of reactions without creating any TICs.

To reduce the solution space and contextualize the model, gene expression data was incorporated into the model. The GPR relation of Arabidopsis as published by Saha *et al* (2011) was consulted to include transcriptomics data from the maturation zone of the roots obtained by Li *et al.* by applying the E-Flux algorithm (Sinha *et al.*, 2021). Metabolic bottleneck analysis identified bottleneck reactions that had formed due to the incorporation of transcriptome datasets into the model. A flow diagram in Supplementary Figure S1 illustrates how MBA is performed on a GEM to identify bottleneck reactions by expanding flux bounds and determining new maximum biomass flux values using FBA.

The associated bottleneck genes were identified from the same Arabidopsis GPR, and gene expression values from published results were used to gain further insights into the nature of the bottleneck genes. The transcriptomics data collected by Geng *et al.* (2013) for the cortex, endodermis, epidermis, and stele root tissue under normal growth conditions and stress-induced growth conditions. The bottleneck genes and their associated gene expression values in each tissue were compared using the t-distributed Stochastic Neighbor Embedding (t-SNE) clustering analysis to evaluate correlations between the expression of bottleneck genes under normal growth conditions and stress-induced growth conditions. The genes associated with each cluster in the t-SNE analysis were further studied using protein-protein interaction networks available in the STRING database (Szklarczyk *et al.*, 2023).

2.3 Biological insights

Dynamic transcriptome datasets collected by Geng *et al.* (2013) were used to analyze flux range variability of the Arabidopsis seedlings (Columbia-0 accession) exposed to media supplemented with 140 mM of sodium chloride (NaCl; salt) (Geng *et al.*, 2013). Spatiotemporal micro-array-based transcriptional profiles were determined for the cortex, endodermis, epidermis, and stele tissues of the root system. The gene expression data collected from the different tissues in normal growth conditions, 1 hour of salt exposure, and 48 hours of salt exposure were incorporated into AraRoot using the E-Flux method to create tissue-specific root models for each of the growth conditions. The normal growth condition was considered as a baseline for the flux ranges while the flux ranges resulting from the 1-hour exposure and 48-hour exposure were compared to the baseline to compare the variations in flux ranges during high salt concentration growth conditions. The categories of flux variation changes that can occur under high salt concentration growth conditions are further illustrated and named in Supplementary Figure S2. The flux ranges found under 1-hour salt exposure were first compared to the flux ranges found under normal conditions, followed by also comparing the 48-hour salt exposure to the normal conditions.

Reduced flux spaces under high salt concentration conditions indicate pathways that produce fewer metabolites under reduced flux, while expanded flux spaces indicate an increase in metabolite production. These conditions are a direct result of the reactions associated with under-expressed and over-expressed genes, respectively. Variations in flux range with no overlap indicate major changes in metabolic reprogramming under stressful

growth conditions. The variations in flux ranges can be confirmed using previously published findings on changes in metabolite production under high salt growth conditions (Kiegle *et al.*, 2000; Geng *et al.*, 2013; Kellermeier *et al.*, 2014; Barberon *et al.*, 2016; Han & Yang, 2021; Smolko *et al.*, 2021).

2.4 Simulation software

All simulations were run on a combination of General Algebraic Modeling System (GAMS) version 24.7.4 and Python version 3.10. In both cases the IBM CPLEX licensed solver was used for optimization calculations such as FBA, MBA, and FVA. All simulations were run on a Linux-based terminal connected to the Holland Computing Center at the University of Nebraska to allow for decreased computational time and adequate memory to run extensive optimization formulations. The final version of AraRoot is available on the SSBio GitHub page (https://github. com/ssbio).

3. RESULTS

3.1 Model initialization

This work presents AraRoot, the first metabolic model that aims to comprehensively capture the biomass formation and stress response of the Arabidopsis root system as described in the literature. The model consists of 2,682 reactions, 2,748 metabolites, and 1,310 genes and was analyzed to determine the metabolic reprogramming that occurs in the root tissues during exposure to stress. Considering Fig. 1, the first step was initializing the model, which involves cross-referencing well-known databases for information on the genome, as well as consulting previously published models of a similar nature. The maize root model published by Chowdhury *et al.* (2022) was considered a good candidate base root model given the major similarities that exist between the protein sequences obtained from maize and Arabidopsis (Saha *et al.*, 2011; Dukowic-Schulze *et al.*, 2014; Schroeder & Saha, 2020a).

To verify the feasibility of using the maize root model as a base model, the genes from the maize and Arabidopsis root systems and their associated protein sequences were included in a BLASTp analysis. Results showed 960 hits between the protein sequences of the two root systems according to the cutoff criteria outlined in Section 2.1, indicating that the metabolic system is likely to be highly similar between both root systems. This justifies the use of the previously published maize root model as a base model for AraRoot. Using the maize root model as a base model, the first step involved updating the composition and stoichiometry of the biomass reaction to reflect the biomass of the Arabidopsis root system rather than that of the maize root system.

Several literature sources were consulted to determine the Arabidopsis root biomass composition. The most noticeable modification to biomass composition in this model includes the addition of cell wall components (cellulose, hemicellulose, and pectin) and the increase in lignin composition, both of which are responsible for structural support and strength in the roots (Zabotina *et al.*, 2012). The specificity of including suberin monomers in the biomass instead of lipids has been well-established in several root studies (Soltis & Soltis, 2021; Pirona *et al.*, 2023). The mass composition of the lignin, pectin, cellulose,

and hemicellulose was obtained from the C3 root composition analyses (Saunders *et al.*, 2006). The specific split between p-hydroxyphenyl lignin, guaiacyl lignin, and syringyl lignin was assumed to be like that of maize.

Hemicellulose as a compound is a complex branched polysaccharide that can vary significantly between plant species (Zabotina *et al.*, 2012; Rao *et al.*, 2023). A study by Zabotina *et al.* (2012) shows that the hemicellulose in Arabidopsis mainly consists of xylan and xyloglucan, with xyloglucan making up most of the hemicellulose composition. Since very little is known about the reaction mechanism of hemicellulose formation, hemicellulose is represented by xylan and xyloglucan in the root model. Saunders *et al.* (2006) were able to quantify the mass composition of the total sugars, fatty acids, amino acids, and suberin monomers found in the root systems of a collection of mostly C3 plants. The sugar compounds and compositions were assumed to be similar to the root system of Arabidopsis proposed by Schroeder and Saha (2020a). The specific weight percentage composition of each fatty acid and suberin monomer was determined by previous experimental studies (Li-Beisson *et al.*, 2013). The specific g/g dry weight compositions of the amino acids in the roots were assumed to be similar to that in the root system of tomatoes, another C3 plant (Gerlin *et al.*, 2022). All Arabidopsis root biomass components and their associated mass percentage compositions can be found in Supplementary File 1, along with a comparison table to show the difference between the original maize root biomass components and that of the final AraRoot model. The cell wall components, pectin and lignin make up most of the root biomass, while the ions involved in biomass formation were reduced compared to the maize root biomass.

Once the biomass reaction in the initial model was updated by changing the stoichiometry and components to match the biomass composition, some reactions had to be added for the model to be able to carry flux through the biomass reaction. The required reactions and their associated mechanisms were found using the KEGG database (Kanehisa *et al.*, 2014). Since all biomass components and their compositions were obtained from the literature, a sensitivity analysis was conducted to quantify the uncertainty of the coefficients of the biomass components in the base model by implementing the algorithm developed by Dinh *et al* (2022). Figure 2A provides the SDR results of the biomass coefficient sensitivity analysis where fluctuations in coefficient values were imposed on three cases: all coefficients at once, only on cellulose, and xylan as singular coefficients. The temporal fluctuations imposed on all coefficients at the same time (c_{ALL}) shows a significant impact on the predicted biomass flux through FBA at an SDR of 14.04%. Imposing a fluctuation on the coefficient of cellulose ($c_{Cellulose}$), a major component in the biomass also results in a significant impact on the biomass predictions with an SDR of 13.96%. The results for imposing fluctuations on the xylan coefficient (c_{Xylan}) shows little impact on biomass formation with an SDR of 1.58%. Expanding on the biomass coefficient sensitivity analysis, the shadow price values of all metabolites in the model having the updated biomass reaction were calculated. To determine the range of shadow price values for the metabolites, the coefficients of the biomass reaction were changed to reflect a 20% fluctuation in the coefficient values. Figure 2B provides the shadow price value ranges for all the biomass metabolites.
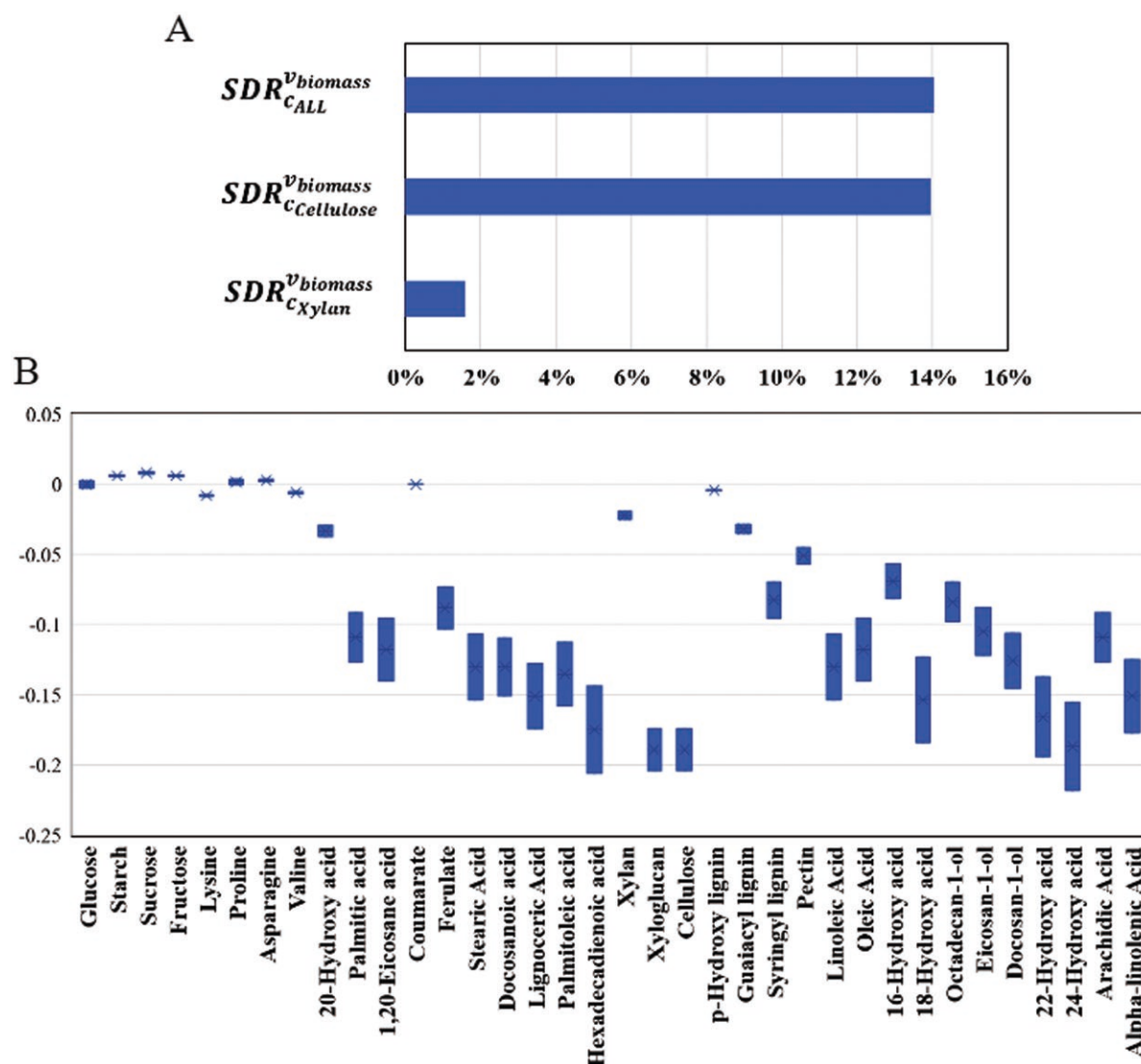
Figure 2. Analysis of the Arabidopsis root biomass composition, including (A) the results of the quantified uncertainty in changes to the biomass coefficients in the model, and (B) the range of shadow price values for the biomass metabolites when coefficient values in the biomass reaction are deviated by 20%.

The positive shadow price values of the sugar components suggest that they act as overflow metabolites, as do the amino acids, proline, and asparagine. All other biomass metabolites were found to have negative shadow price values, indicating that they act as growth-limiting metabolites. The shadow price values calculated for the suberin-derived fatty acids, hexadecadienoic acid, and 22-hydroxydocosanoic acid, and the cell wall-derived components xyloglucan and cellulose indicate that these metabolites would have the highest correlation to biomass synthesis in the roots. The suberin-related metabolites make up a minute part of the biomass composition, whereas cellulose and xyloglucan within the cell wall make up relatively large portions of the biomass composition.

3.2 Model reconstruction

3.2.1 *Resolved blocked reactions and TICs.*

At this stage, the base model included 4,089 reactions, 4,419 metabolites, and 6,389 genes. Conducting FVA on the model identifies blocked reactions by identifying all the reactions that have a maximum and minimum optimal flux of zero when producing biomass in the case of AraRoot. The base model was found to have 2,315 blocked reactions. To determine which of the blocked reactions needed to remain in the model due to possible downstream involvement in different pathways, the list of blocked reactions was compared to the reactions that are associated with genes that have high homology between Arabidopsis and maize as obtained from the BLAST results. From the comparison, 678 reactions were found to have high homology between Arabidopsis and maize and would need to remain in the model, despite not carrying flux during biomass production. The remaining 1,637 blocked reactions were then manually scrutinized to ensure that downstream reactions were also kept in the model. Of the 1,637 blocked reactions, 22 reactions were found to play a crucial role in downstream reactions and were kept in the model. A total of 1,615 blocked reactions were finally removed from the model by deleting the reactions and removing any unused associated metabolites.
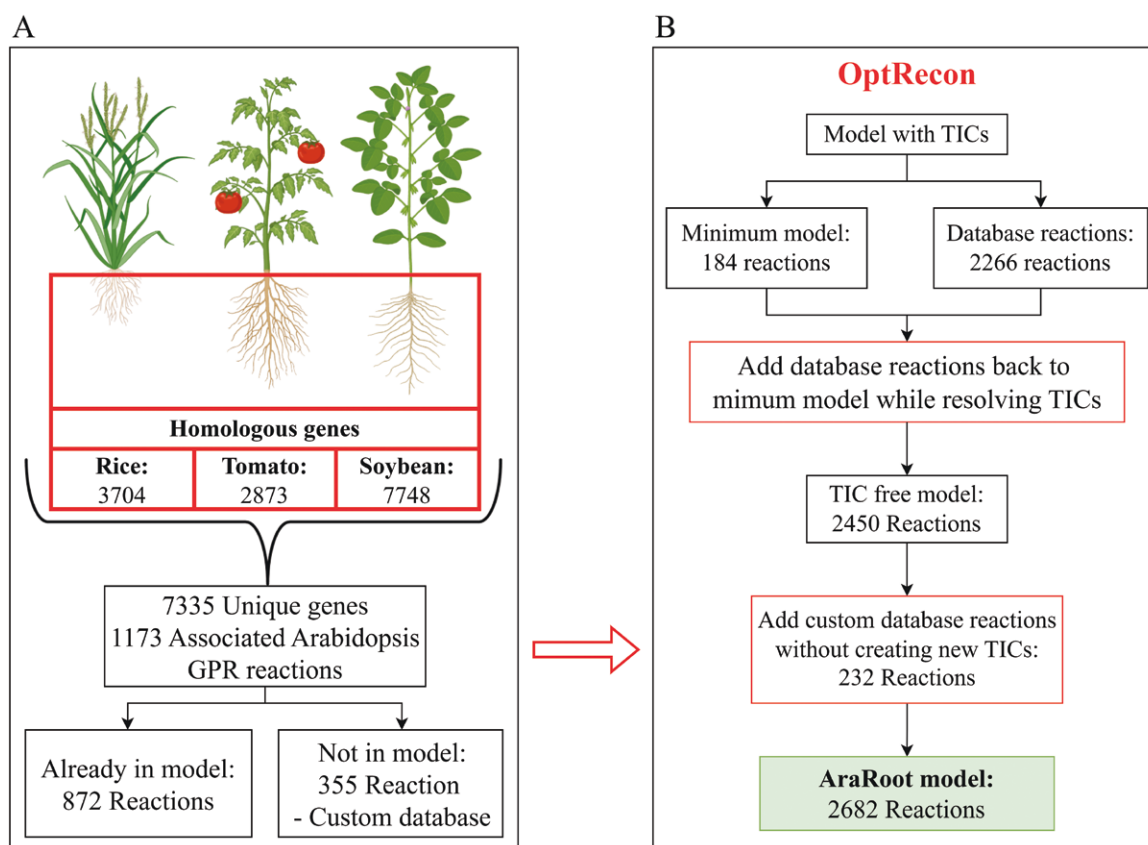
Figure 3. Resolving missing reactions in the model by (A) creating a custom database from other C3 plants, followed by (B) using OptRecon to add reactions to the model from a custom database without creating any TICs.

Next, all missing specialized metabolites were verified to be in the Arabidopsis roots and their associated missing reactions were added to the model (Strehmel *et al.*, 2014). Specialized metabolites that were added include coniferin, syringin, lariciresinol, esculetin, scopoletin, esculin, and scopoline (Strehmel *et al.*, 2014). Once the specialized metabolites and reactions were added to the model it could be evaluated for any TICs. Since a previously published maize root model was used as the base model for AraRoot the inherent base model did not contain any TICs, as confirmed by the tool, OptRecon. Once this was confirmed, it was necessary to consider that the model might be missing some reactions that could be present in Arabidopsis but would be missing from the maize root model.

To find and add any missing reactions, a custom database containing reactions from the Arabidopsis genome was created. The SBML file containing all the custom database reactions can be found in the AraRoot GitHub repository (https://github.com/ssbio/AraRoot), and the compilation of the genes used to create the custom database can be found in Supplemental File 4. These reactions were identified by obtaining transcriptome data of the root systems of other C3 plants, specifically rice (Liu *et al.*, 2021), tomato (Gerlin *et al.*, 2022), and soybean (Adhikari *et al.*, 2019). The genes that are expressed in these root systems were found to have homologous genes that are expressed in Arabidopsis roots. Referring to the GPR relationship of Arabidopsis, a list of reactions was identified that are associated with genes that have

high homology among C3 root systems. These reactions were considered to form the custom database from which OptRecon could add reactions to account for missing reactions. Figure 3 illustrates the steps taken to create the custom database and use OptRecon to fill the missing reactions into the model, constructing the final AraRoot model.

A list of 7,335 unique Arabidopsis genes were found to have high homology with the root systems of rice, tomato, and soybean. From these genes, 355 reactions were identified from the published *A. thaliana* iRS1597 (Saha *et al.*, 2011) model GPR and were used as a custom reaction database for OptRecon. From the custom database, 232 reactions were added to the model. The final AraRoot model consists of 2,682 reactions, 2,748 metabolites, and 1,310 genes.

3.2.2 *The inclusion of transcriptome datasets.*

At this stage the bounds of the fluxes in the model were all set to large values, essentially 'unbounding' the flux ranges. When computing FBA with unbounded flux ranges, the possible solutions to optimizing the maximum biomass flux ($v_{biomass}$) are infinite and can take up very large unrealistic fluxes through the reaction pathways in the model. The best practice to reduce the solution space and contextualize the model is to include empirical transcriptomic data of the Arabidopsis root system in the model. The constricted flux bounds will also result in more realistic fluxes through the metabolic pathways as compared to what happens in nature.

Gene expression data obtained from literature was incorporated into the model to contextualize the reaction bounds to more realistic limits. The work by Li *et al.* (2016) generated a high-resolution spatiotemporal map of gene expression in the Arabidopsis root system. The fragments per kilobase per million mapped fragments (FPKM) values of 23,934 unique genes were obtained for the meristematic, elongation, and maturation zones of the root system. From the Arabidopsis GPR, 1,310 unique genes are associated with 1,171 reactions from the whole-plant Arabidopsis genome. Of the 1,171 gene-associated reactions, 507 were present in the AraRoot model and the associated bounds were adjusted using the E-Flux algorithm.

3.2.3 MBA results.

Following the procedure for MBA, 106 bottleneck reactions were identified, which correlate with 158 unique Arabidopsis genes according to the GPR, of which 145 genes were found to have high gene expression values under both normal growth conditions and high salt concentration growth conditions. Transcriptomic data collected by Geng *et al.* (2013) included gene expression values of four different tissues in the root under normal growth conditions and high salt growth conditions. The tissue-specific transcriptomic data for all genes was included in the model using E-Flux. t-Distributed Stochastic Neighbor Embedding (t-SNE) clustering analysis revealed further insights into the correlations in expression between the proposed bottleneck genes in normal and salt conditions. Figure 4 provides the clustering results of all bottleneck genes for each cell type in the root to determine the metabolic correlations between the bottleneck genes when the Arabidopsis roots are exposed to high salt concentration growth conditions as compared to normal salt conditions. Supplementary File 5 provides a list of all the bottleneck genes, the associated gene expression values, the corresponding AraRoot reactions according to the GPR, and the cluster and pathway that each gene is associated with as illustrated in Fig. 4.

Based on visualization, in the endodermis and stele, the bottleneck genes were found to form distinct clusters where the majority of genes in both cell types are associated with the oxidative phosphorylation, and phosphonate and phosphinate metabolism pathways. Phosphonate and phosphinate metabolism are also prevalent in the cortex and epidermis, though the clusters are not as distinct as in the endodermis and stele. Several bottleneck genes were found to impact the TCA cycle and pyrimidine metabolism in the cortex and epidermis in high salt concentration growth conditions. The bottleneck genes impacting the ubiquinone and other terpenoid-quinone biosynthesis pathways in the endodermis and epidermis formed a distinct cluster in each case, indicating a high correlation during exposure to normal and salt conditions. The cortex seems to be the only cell type with bottleneck genes involved in the lysine degradation pathway, which is a downstream pathway of the TCA cycle. The bottleneck genes and their associated gene expression values in each identified cluster were imported into the STRING database and the protein-protein interaction networks found for Arabidopsis provided further insights into the correlations between the bottleneck genes in each root tissue under salt stress growth conditions. The illustrative outputs of the protein-protein interaction networks can be found in Supplementary File 3. From comparing the protein-protein networks of each tissue to one another it was found that some proteins found in each of the four networks include PBL1, ALA2, PECT1, and CCT1.

3.3 Biological insights

To better understand the capabilities AraRoot has for capturing the metabolic reprogramming that occurs in stress conditions, a flux range variation analysis was carried out to investigate the changes in reaction fluxes in the roots when exposed to high salt concentration growth conditions. The transcriptomic data published by Geng *et al.* (2013) included gene expression values of the cortex, endodermis, epidermis, and stele tissues in the root system that were exposed to normal growth conditions, 1 hour of salt exposure (short exposure), and 48 hours of salt exposure (long exposure). This data was implemented into AraRoot through the E-Flux algorithm to formulate 12 different tissue-specific models under various growth conditions. The flux ranges found under normal growth conditions were considered the baseline fluxes, and the flux ranges that resulted from the high salt concentration growth conditions were each compared to the normal growth condition flux ranges according to the categories described in Figure S2 in Supplemental File 3. The resulting variations in flux range from normal growth conditions to high salt concentration growth conditions are provided in Fig. 5.

Central carbon metabolism includes the glycolysis, pentose phosphate, and TCA cycle pathways, which are responsible for the conversion of carbon-containing compounds to sugars and organic molecules and are perhaps the best starting point when analyzing the metabolic reprogramming of the roots under salt stress (Dieuaide-Noubhani *et al.*, 2007; Nägele *et al.*, 2010). Other major pathways that appear upstream and downstream of central carbon metabolism include the *de novo* fatty acid synthesis, galactose, purine, and glycerophospholipid pathways.

Most of the pathways in the cortex experience a shrunk flux space after 1 hour of salt exposure, but prolonged salt exposure led to widened flux spaces in all pathways except the pentose phosphate pathway. The endodermis and stele seem to experience the opposite metabolic reprogramming, experiencing widened flux spaces after 1 hour of salt exposure while prolonged salt exposure leads to shrunk flux spaces in all pathways except the pentose phosphate pathway. The epidermis seems to experience overall shrunk flux spaces in both short and long salt exposure times.

4. DISCUSSION

4.1 Reconstructing AraRoot

This work introduces AraRoot, a metabolic model of the Arabidopsis root system that includes a comprehensive biomass composition that for the first time has the capability of capturing the metabolic reprogramming that occurs in separate root tissues when exposed to stress-induced growth conditions. The development of a tissue-specific metabolic model such as AraRoot is crucial for gaining insights into the unique metabolic functions and adaptations of distinct plant organs, such as the roots. AraRoot successfully allows for a more precise understanding
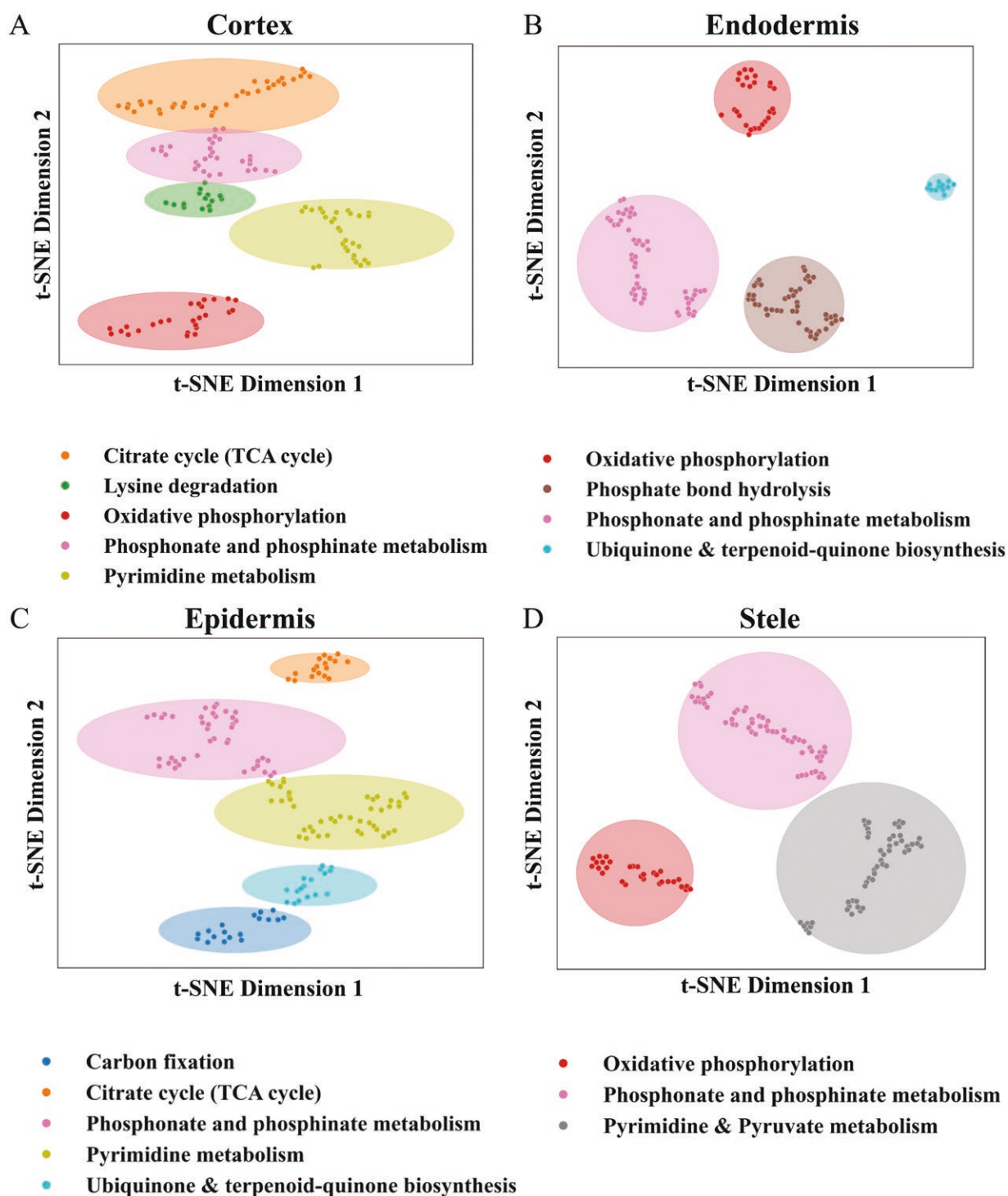
Figure 4. K-Means clustering analysis of bottleneck genes in the (A) Cortex, (B) Endodermis, (C) Epidermis, and (D) Stele to indicate genetic correlations of the bottleneck genes in normal and salt conditions.

of how specific root tissues, including the cortex, endodermis, epidermis, and stele, respond to environmental stimuli, which advances our knowledge of stress responses in plant biology and can be further expanded to other crop GEMs to gain insights on how to improve crop resilience.

Our in-house tool, OptRecon, was instrumental in the development of AraRoot. OptRecon facilitated the accurate reconstruction of the Arabidopsis root metabolic network by removing TICs and ensuring the completeness of the model by adding missing reactions from a database. This tool enabled us to refine the metabolic pathways specific to the different root tissues, thereby enhancing the predictive power and reliability of AraRoot. The successful application of OptRecon in this study underscores its potential as a valuable resource for developing other tissue-specific GEMs, further contributing to the field of plant systems biology.
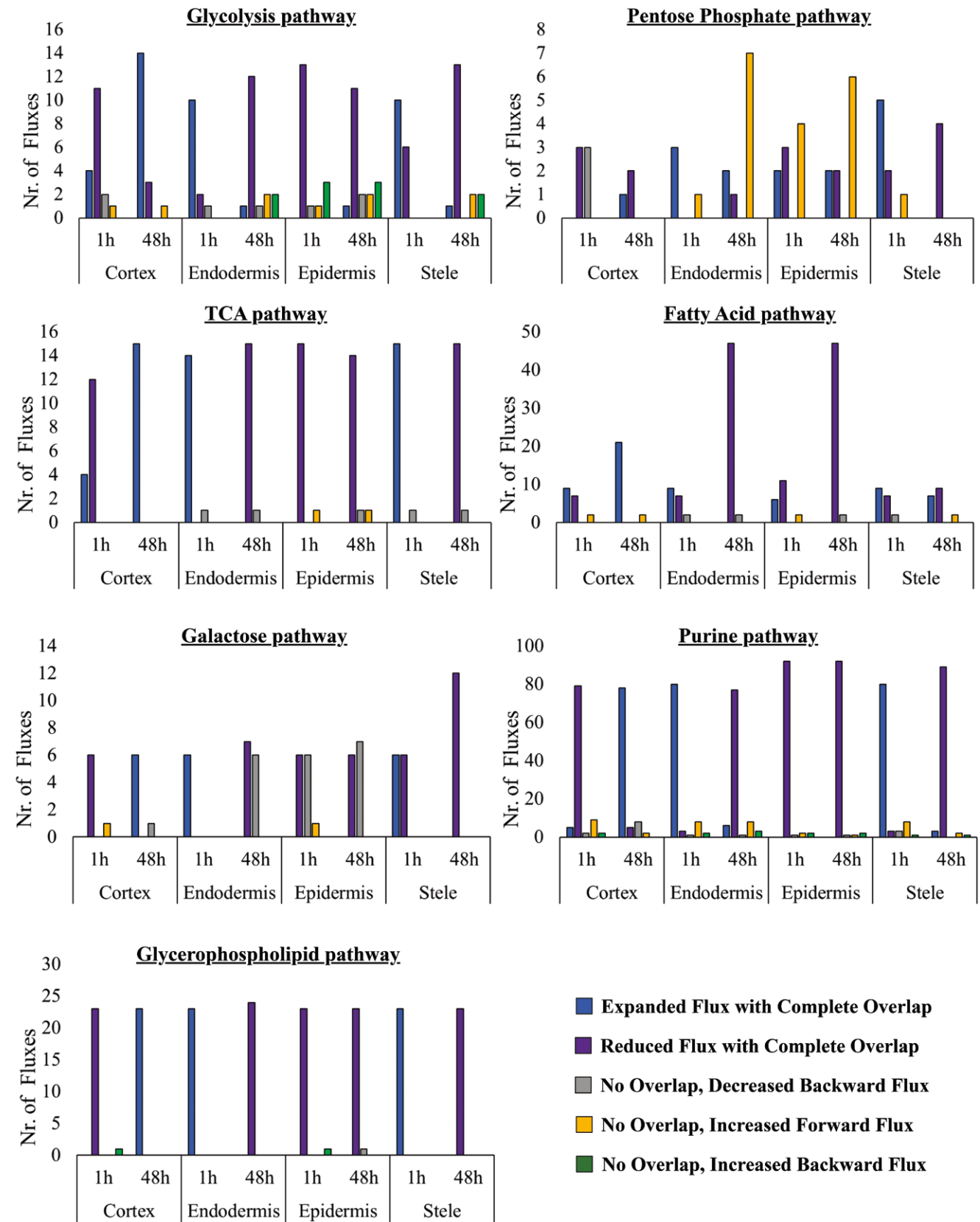
Figure 5. Flux range variation resulting from high salt exposure during root growth.

Another crucial part of developing AraRoot was ensuring that the model comprehensively captures a realistic biomass composition of the root organ as described in multiple literature sources (Zabotina *et al.*, 2012; Li-Beisson *et al.*, 2013; Schroeder & Saha, 2020a; Gerlin *et al.*, 2022; Chowdhury *et al.*, 2023). A comprehensive biomass composition would provide a realistic representation of the metabolic demands and outputs of the roots, leading to more accurate simulations and predictions of metabolic behavior under various conditions. The integration of such precise biomass data ensures that the model can better replicate the physiological and biochemical states of tissues under different growth conditions. The biomass composition included in AraRoot was determined from several literature sources and the sensitivity analysis of the model to changes in the biomass composition and stoichiometry highlights the need to experimentally verify the actual biomass composition of Arabidopsis roots. The shadow price values of the biomass components indicate that most of the sugar components act as overflow metabolites. When considering the starch and sucrose metabolism pathways, glucose is formed from D-glucose 1-phosphate, which is also needed to form the cellulose and hemicellulose components of the cell wall. Since cellulose makes up most of the biomass composition, higher flux is required through the cellulose formation reaction within the model, resulting in glucose being an overflow metabolite. The negative shadow price values indicate growth-limiting metabolites. From the shadow price results, cellulose and xyloglucan within the cell wall would impact the actual biomass composition more than any of the other biomass metabolites since they have highly negative shadow price values and make up a relatively large proportion of the biomass composition. The negative shadow price value range of −0.18 to 0.2 of cellulose ties in with the sensitivity analysis of the model after perturbations in the cellulose coefficients resulting in an SDR of 14%, both indicating that changes to the amount of cellulose in the biomass composition could highly affect the model outcomes. The shadow price value of xylan was found to be close to zero, while also indicating that it is a growth-limiting metabolite. This corresponds with the results of the sensitivity analysis with an SDR of 1.58%, both indicating that changes in the xylan composition of the biomass would not have a significant effect on the biomass predictions by AraRoot.

Additionally, the fatty acids, suberin monomers, and lignin components of root biomass all exhibit growth-limiting properties. Both the fatty acids and the suberin monomers depend on acetyl-CoA as their major precursor metabolite and are dependent on flux through the de novo fatty acid pathway for producibility. Although the fatty acids and suberin monomers make up a smaller proportion of the biomass, perturbations in the coefficients of the components affect the other components in the groups as well, creating a 'ripple effect' in the biomass production predictions, making these metabolites growth-limiting metabolites. The insights gained from analyzing the bottleneck genes identified from the model provide a unique perspective on the metabolic behavior of the root system and would have taken much longer to gain from empirical investigations, highlighting the incredible power of utilizing tissue-specific predictive metabolic models.

4.2 Predicting stress responses

Developing the AraRoot metabolic model before contextualizing it to represent specific root cell types provided a foundational framework that encompasses the comprehensive metabolic capabilities of the root. This includes all known metabolic reactions and pathways within the root, providing a holistic understanding that can be selectively refined and tailored to reflect the metabolic activities of specific cell types. Furthermore, the development of a whole-root metabolic model allows for the utilization of a wide range of experimental data, including transcriptomics, proteomics, and metabolomics data, and allows for easy adaptation or expansion as new omics data becomes available. To analyze the capability of AraRoot in capturing the response of the root system to stress, the model was contextualized to represent the cortex, endodermis, epidermis, and stele tissues in the root system. Publicly available transcriptomic data for these tissues were utilized to not only contextualize the model but also to analyze the nature of the bottleneck genes identified from the model.

From the clustering analysis results of the transcriptome data of the bottleneck genes, it can be concluded that phosphorus-containing pathways and energy-related pathways seem to be consistently involved in bottlenecks when the Arabidopsis roots are exposed to high salt stress conditions. The bottleneck genes were found to be associated with the phosphonate and phosphinate metabolism pathway in each of the root tissues. By further analyzing the protein-protein networks of the clusters it was found that a common bottleneck protein among all cell types is PBL1, which is a receptor-like cytoplasmic kinase responsible for signaling with calcium and inhibiting root growth during stress conditions. Experimental work published by Kiegle *et al.* (2000) validated that calcium is vital for signaling responses in the Arabidopsis root system when exposed to salt stress. Calcium signaling was also found to be cell type-specific in the roots (Kiegle *et al.*, 2000). Other bottleneck proteins found to be associated with all four root tissue types include ALA2, PECT1, and CCT1, all of which are associated with phospholipid biosynthesis.

To better understand the flux range variation results, it was necessary to consult previous experimental findings. Several studies have shown that during the initial stages of root exposure to salt conditions, the roots go into a 'shocked' state where growth and metabolism are halted. However, after several hours of exposure, the roots return to a somewhat normal rate of growth (Iyer-Pascuzzi *et al.*, 2011; Machado & Herrgård, 2014; Pirona *et al.*, 2023). From the flux variability results in Fig. 5, this is observed especially in the cortex, where all pathways have shrunken flux spaces after 1 hour of salt exposure but have widened flux spaces after 48 hours of salt exposure. However, in the pentose phosphate pathway, the formation of D-ribulose 5-phosphate and D-glyceraldehyde 3-phosphate remains suppressed in the cortex under high salt growth conditions. The de novo fatty acid, galactose, and purine pathways undergo the same reprogramming as the glycolysis pathway within the cortex.

The endodermis seems to undergo the opposite reprogramming as compared to the cortex. After 1 hour of salt exposure, the majority of the pathways and their associated reactions experience an increase in flux a widened flux space, and a shrunken

flux space after 48 hours of exposure, with the exception of the pentose phosphate pathway. It was also observed by Kiegle *et al.* (2000) that the endodermis has a unique response to salt exposure, with higher levels of cytoplasmic free calcium for stress signaling (Kiegle *et al.*, 2000). The endodermis is also responsible for transporting water from the outer cortex and epidermis to the stele, which transports the water, via the xylem, to the rest of the plant (Brady *et al.*, 2007; Dinneny *et al.*, 2008). A separate study showed that the endodermis enhances the deposition of suberin in response to long-term exposure to salt stress, likely to limit water loss from the root (Barberon *et al.*, 2016). Reduced metabolic activity in the endodermis after 48 hours of salt exposure could be due to the suppression of water uptake from the growth media.

The increased flux range in the endodermis experienced by the pentose phosphate pathway is due to an upregulation in the formation of D-ribulose 5-phosphate. A study published by Huang *et al.* (2020) confirmed that D-ribulose 5-phosphate forms an essential part of the cell wall biosynthesis pathway as one of the essential compounds needed to form cell walls and seems to be upregulated in both the endodermis and epidermis during growth in high salt conditions (Huang *et al.*, 2020). The upregulation of cell wall biosynthesis would allow for cells to maintain internal water retention during high salt growth conditions and provide better support for the root structure (Kiegle *et al.*, 2000; Iyer-Pascuzzi *et al.*, 2011; Zabotina *et al.*, 2012).

Unlike the other cell types, the epidermis was found to have similar metabolic behavior after both 1 hour and 48 hours of high salt growth conditions. All pathways, except the pentose phosphate pathway, were found to have reduced flux spaces in response to salt conditions. However, like the endodermis, the formation of D-ribulose 5-phosphate was upregulated in the pentose phosphate pathway after both 1 and 48 hours of salt exposure. The study published by Dinneny *et al.* (2008) confirms that the epidermis is the only cell type in the roots for which gene expression is highly responsive to environmental conditions (Dinneny *et al.*, 2008). The metabolic reprogramming by the stele cells under high salt conditions is very similar to that of the endodermis. Considering that the stele contains both the xylem and the phloem, which are responsible for transporting water and nutrients to the rest of the plant, it can be expected that metabolic activity will be suppressed under prolonged exposure to high salt conditions and result in reduced flux spaces in all pathways after 48 hours of salt exposure. Results also indicate that the glycerophospholipid pathway undergoes complete reprogramming in the cortex, endodermis, and stele but remains suppressed in the epidermis during short and long exposures to salt (Han & Yang, 2021). The expanded flux ranges in the glycerophospholipid pathway in the cortex also indicate that signaling during salt stress growth might be regulated in the cortex.

Considering the overall impact of high salt growth conditions, the different tissue types of the roots undergo unique metabolic reprogramming to ensure that the plant survives. The cortex undergoes an immediate metabolic suppression upon initial exposure to salt conditions; however, it is the tissue most likely responsible for root growth, as evidenced by increased synthesis of sugar and organic compounds after prolonged exposure to salt conditions. The metabolic activity in the endodermis, epidermis, and stele is suppressed after prolonged salt exposure, to possibly

provide a counter-pressure to the external salt conditions while maintaining internal functionality. The pentose phosphate pathway in the endodermis and epidermis immediately starts producing more D-ribulose 5-phosphate that aids in cell wall biosynthesis to allow for thicker cell walls, which may lead to better water retention in the cells and thickened layers of protection for the roots themselves. Comparing the flux range variation results of AraRoot to previous experimental findings, these results show that AraRoot can accurately capture the metabolic reprogramming that occurs in the root tissues during growth in high salt concentrations as described in literature (Kiegle *et al.*, 2000; Dinneny *et al.*, 2008; Iyer-Pascuzzi *et al.*, 2011; Zabotina *et al.*, 2012; Geng *et al.*, 2013; Smolko *et al.*, 2021).

5. CONCLUSION

The development of the root tissue-specific metabolic model, AraRoot, highlights the importance of metabolic models in studying the metabolic behavior of plants. Not only does the model capture the comprehensive biomass formation of the Arabidopsis root system, but it also succeeds in providing valuable insights into the metabolic reprogramming that occurs in different root tissues during exposure to high salt concentrations. Our work highlights the importance of the cortex to root growth in stress conditions, while also capturing the role of the endodermis and epidermis in producing thicker cell walls to allow for better water retention. AraRoot could be used to further predict the metabolic responses of the root system to other stress conditions, such as temperature, pH, drought, and nutrient starvation. The methodology proposed here could also be used to develop metabolic models for the root systems of major crops. This would allow for further insights into how crops could be improved in terms of resilience to environmental stress and yield.

SUPPLEMENTARY DATA

Supplementary data is available at in silico Plants online. The following additional information is available in the online version of this article.

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

L.E.: Conceptualized problem, reconstructed model, performed analyses, analyzed results, and wrote the original draft. N.A.: Performed analyses, aided in analyzing results, aided in writing the manuscript. M.D.Y.-N., B.J.N., and E.E.S.: Provided insights

into the metabolism of Arabidopsis, reviewed and edited manuscript R.S.: Procured funding, supervised, provided insights into the methodology, reviewed and edited the manuscript.

CONFLICT OF INTEREST

Nothing to declare.

DATA AVAILABILITY

All models and codes used in this work can be found in the following GitHub directory: (https://github.com/ssbio/AraRoot).

REFERENCES

Adhikari S, Damodaran S, Subramanian S. Lateral root and nodule transcriptomes of soybean. *Data* 2019;4:64. https://doi.org/10.3390/data4020064

Altschul SF, Gish W, Miller W et al. Basic local alignment search tool. *J Mol Biol* 1990;215:403–10. https://doi.org/10.1016/S0022-2836(05)80360-2

The Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature* 2000;408:796–815. https://doi.org/10.1038/35048692

Barberon M, Vermeer JEM, De Bellis D et al. Adaptation of root function by nutrient-induced plasticity of endodermal differentiation. *Cell* 2016;164:446–59. https://doi.org/10.1016/j.cell.2015.12.021

Becker SA, Palsson BO. Context-specific metabolic networks are consistent with experiments. *PLoS Comput Biol* 2008;4:e1000082. https://doi.org/10.1371/journal.pcbi.1000082

Berardini TZ, Reiser L, Li D et al. The Arabidopsis information resource: making and mining the 'gold standard' annotated reference plant genome. *Genesis* 2015;53:474–85. https://doi.org/10.1002/dvg.22877

Brady SM, Orlando DA, Lee JY et al. A high-resolution root spatiotemporal map reveals dominant expression patterns. *Science* 2007;318:801–6. https://doi.org/10.1126/science.1146265

Brendel, V., Kurtz, S., & Walbot, V. Comparative genomics of Arabidopsis and maize: prospects and limitations. *Genome Biol* 2002;3:1. https://doi.org/10.1186/gb-2002-3-3-reviews1005

Chen ZJ, Wang J, Tian L et al.; others. The development of an Arabidopsis model system for genome-wide analysis of polyploidy effects. *Biol J Linn Soc* 2004;82:689–700. https://doi.org/10.1111/j.1095-8312.2004.00351.x

Chowdhury NB, Schroeder WL, Sarkar D et al. Dissecting the metabolic reprogramming of maize root under nitrogen-deficient stress conditions. *J Exp Bot* 2022;73:275–91. https://doi.org/10.1093/jxb/erab435

Chowdhury NB, Simons-Senftle M, Decouard B et al. A multi-organ maize metabolic model connects temperature stress with energy production and reducing power generation. *IScience* 2023;26:108400. https://doi.org/10.1016/j.isci.2023.108400

Colijn C, Brandes A, Zucker J et al. Interpreting expression data with metabolic flux models: predicting mycobacterium tuberculosis mycolic acid production. *PLoS Comput Biol* 2009;5:e1000489. https://doi.org/10.1371/journal.pcbi.1000489

Dal'Molin C. G. de O., Quek LE, Palfreyman RW et al. AraGEM, a genome-scale reconstruction of the primary metabolic network in Arabidopsis. *Plant Physiol* 2010;152:579–89. https://doi.org/10.1104/PP.109.148817

Dal'Molin CGO, Quek LE, Palfreyman RW et al. Plant genome-scale modeling and implementation. *Methods Mol Biol* 2014;1090:317–32. https://doi.org/10.1007/978-1-62703-688-7_19

de Oliveira Dal'Molin CG, Quek LE, Saa PA et al. A multi-tissue genome-scale metabolic modeling framework for the analysis of whole plant systems. *Front Plant Sci* 2015;6:1–12. https://doi.org/10.3389/fpls.2015.00004

Dieuaide-Noubhani M, Alonso A-P, Rolin D et al. Metabolic flux analysis: recent advances in carbon metabolism in plants. *Plant Syst Biol* 2007;97:213–43. https://doi.org/10.1007/978-3-7643-7439-6_10

Dinh HV, Sarkar D, Maranas CD. Quantifying the propagation of parametric uncertainty on flux balance analysis. *Metab Eng* 2022;69:26–39. https://doi.org/10.1016/j.ymben.2021.10.012

Dinneny JR, Long TA, Wang JY et al. Cell identity mediates the response of Arabidopsis roots to abiotic stress. *Science* 2008;320:942–5. https://doi.org/10.1126/science.1153795

Dukowic-Schulze S, Harris A, Li J et al. Comparative Transcriptomics of early meiosis in Arabidopsis and maize. *J Genet Genomics* 2014;41:139–52. https://doi.org/10.1016/j.jgg.2013.11.007

Edwards JS, Palsson BO. Systems properties of the Haemophilus influenzae Rd metabolic genotype. *J Biol Chem* 1999;274:17410–6. https://doi.org/10.1074/jbc.274.25.17410

Geng Y, Wu R, Wee CW et al. A spatio-temporal understanding of growth regulation during the salt stress response in Arabidopsis. *Plant Cell* 2013;25:2132–54. https://doi.org/10.1105/tpc.113.112896

Gerlin L, Cottret L, Escourrou A et al. A multi-organ metabolic model of tomato predicts plant responses to nutritional and genetic perturbations. *Plant Physiol* 2022;188:1709–23. https://doi.org/10.1093/plphys/kiab548

Gudmundsson S, Thiele I. Computationally efficient flux variability analysis. *BMC Bioinf* 2010;11:1–3. https://doi.org/10.1186/1471-2105-11-489

Han, X., & Yang, Y. Phospholipids in salt stress response. *Plants* 2021;10:1. https://doi.org/10.3390/plants10102204

Huang J-B, Zou Y, Zhang X et al. RIBOSE PHOSPHATE ISOMERSASE 1 influences root development by acting on cell wall biosynthesis, actin organization, and auxin transport in Arabidopsis. *Front Plant Sci* 2020;10:1641. https://doi.org/10.3389/fpls.2019.01641

Iyer-Pascuzzi AS, Jackson T, Cui H et al. Cell identity regulators link development and stress responses in the Arabidopsis root. *Dev Cell* 2011;21:770–82. https://doi.org/10.1016/j.devcel.2011.09.009

Kanehisa M, Goto S, Sato Y et al. Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res* 2014;42:D199–205. https://doi.org/10.1093/nar/gkt1076

Kellermeier F, Armengaud P, Seditas TJ et al. Analysis of the root system architecture of Arabidopsis provides a quantitative readout of crosstalk between nutritional signals. *Plant Cell* 2014;26:1480–96. https://doi.org/10.1105/tpc.113.122101

Kiegle E, Moore CA, Haseloff J et al. Cell-type-specific calcium responses to drought, salt and cold in the Arabidopsis root. *Plant J* 2000;23:267–78. https://doi.org/10.1046/j.1365-313x.2000.00786.x

Li S, Yamada M, Han X et al. High-resolution expression map of the Arabidopsis root reveals alternative splicing and lincRNA regulation. *Dev Cell* 2016;39:508–22. https://doi.org/10.1016/j.devcel.2016.10.012

Li-Beisson Y, Shorrosh B, Beisson F et al. Acyl-lipid metabolism. *Arabidopsis Book* 2013;11:e0161. https://doi.org/10.1199/tab.0161

Liu Q, Liang Z, Feng D et al. Transcriptional landscape of rice roots at the single-cell resolution. *Mol Plant* 2021;14:384–94. https://doi.org/10.1016/j.molp.2020.12.014

MacGillivray M, Ko A, Gruber E et al. Robust analysis of fluxes in genome-scale metabolic pathways. *Sci Rep* 2017;7:268. https://doi.org/10.1038/s41598-017-00170-3

Machado D, Herrgård M. Systematic evaluation of methods for integration of transcriptomic data into constraint-based models of metabolism. *PLoS Comput Biol* 2014;10:e1003580. https://doi.org/10.1371/journal.pcbi.1003580

Mahadevan R, Edwards JS, Doyle FJ. Dynamic flux balance analysis of diauxic growth in Escherichia coli. *Biophys J* 2002;83:1331–40. https://doi.org/10.1016/S0006-3495(02)73903-9

Metcalf AJ, Nagygyor A, Boyle NR. Rapid Annotation of Photosynthetic Systems (RAPS): automated algorithm to generate genome-scale metabolic networks from algal genomes. *Algal Res* 2020;50:101967. https://doi.org/10.1016/j.algal.2020.101967

Mintz-Oron S, Meir S, Malitsky S et al. Reconstruction of Arabidopsis metabolic network models accounting for subcellular compartmentalization and tissue-specificity. *Proc Natl Acad Sci USA* 2012;109:339–44. https://doi.org/10.1073/pnas.1100358109

Nägele T, Henkel S, Hörmiller I et al. Mathematical modeling of the central carbohydrate metabolism in Arabidopsis reveals a substantial regulatory influence of vacuolar invertase on whole plant carbon metabolism. *Plant Physiol* 2010;153:260–72. https://doi.org/10.1104/pp.110.154443

Nelson E, Saha R. *Automated genome-scale model reconstruction upon resolving thermodynamically infeasible cycles*, 2024.

Ngo AH, Nakamura Y. Phosphate starvation-inducible GLYCERO-PHOSPHODIESTER PHOSPHODIESTERASE6 is involved in Arabidopsis root growth. *J Exp Bot* 2022;73:2995–3003. https://doi.org/10.1093/jxb/erac064

Pirona R, Frugis G, Locatelli F et al. Transcriptomic analysis reveals the gene regulatory networks involved in leaf and root response to osmotic stress in tomato. *Front Plant Sci* 2023;14:1155797. https://doi.org/10.3389/fpls.2023.1155797

Poolman MG, Miguet L, Sweetlove LJ et al. A genome-scale metabolic model of Arabidopsis and some of its properties. *Plant Physiol* 2009;151:1570–81. https://doi.org/10.1104/pp.109.141267

Provart NJ, Alonso J, Assmann SM et al.; others. 50 years of Arabidopsis research: highlights and future directions. *New Phytol* 2016;209:921–44. https://doi.org/10.1111/nph.13687

Rao J, Lv Z, Chen G et al. Hemicellulose: structure, chemical modification, and application. *Prog Polym Sci* 2023;140:101675. https://doi.org/10.1016/j.progpolymsci.2023.101675

Saha R, Suthers PF, Maranas CD. Zea mays iRS1563: a comprehensive genome-scale metabolic reconstruction of maize metabolism. *PLoS One* 2011;6:e21784. https://doi.org/10.1371/journal.pone.0021784

Saunders CJ, Megonigal JP, Reynolds JF. Comparison of belowground biomass in C 3-and C 4-dominated mixed communities in a Chesapeake Bay brackish marsh. *Plant Soil* 2006;280:305–22. https://doi.org/10.1007/s11104-005-3275-3

Schroeder WL, Saha R. Introducing an optimization- and explicit Runge-Kutta- based approach to perform dynamic flux balance analysis. *Sci Rep* 2020a;10:9241. https://doi.org/10.1038/s41598-020-65457-4

Schroeder WL, Saha R. OptFill: a tool for infeasible cycle-free gapfilling of stoichiometric metabolic models. *IScience* 2020b;23:100783. https://doi.org/10.1016/j.isci.2019.100783

Shaw R, Cheung CYM. A dynamic multi-tissue flux balance model captures carbon and nitrogen metabolism and optimal resource partitioning during arabidopsis growth. *Front Plant Sci* 2018;9:324324. https://doi.org/10.3389/FPLS.2018.00884

Sinha N, van Schothorst EM, Hooiveld GJEJ et al. Exploring the associations between transcript levels and fluxes in constraint-based models of metabolism. *BMC Bioinf* 2021;22:1. https://doi.org/10.1186/s12859-021-04488-8

Smolko A, Bauer N, Pavlović I et al. Altered root growth, auxin metabolism and distribution in Arabidopsis thaliana exposed to salt and osmotic stress. *Int J Mol Sci* 2021;22:7993. https://doi.org/10.3390/ijms22157993

Soltis PS, Soltis DE. Plant genomes: markers of evolutionary history and drivers of evolutionary change. *Plants People Planet* 2021;3:1. https://doi.org/10.1002/ppp3.10159

Strehmel N, Böttcher C, Schmidt S et al. Profiling of secondary metabolites in root exudates of Arabidopsis thaliana. *Phytochem* 2014;108:35–46. https://doi.org/10.1016/j.phytochem.2014.10.003

Szklarczyk D, Kirsch R, Koutrouli M et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res* 2023;51:1. https://doi.org/10.1093/nar/gkac1000

Thiele I, Palsson B. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat Protocols* 2010;5:93–121. https://doi.org/10.1038/nprot.2009.203

Varma A, Palsson BO. Metabolic flux balancing: basic concepts, scientific and practical use. *Bio/Technology* 1994;12:994–8. https://doi.org/10.1038/nbt1094-994

Woodhouse MR, Cannon EK, Portwood JL et al. A pan-genomic approach to genome databases using maize as a model system. *BMC Plant Biol* 2021;21:1–10. https://doi.org/10.1186/s12870-021-03173-5

Xu H, Liu P, Wang C et al. Transcriptional networks regulating suberin and lignin in endodermis link development and ABA response. *Plant Physiol* 2022;190:1165–81. https://doi.org/10.1093/plphys/kiac298

Xu W, Ding G, Yokawa K et al. An improved agar-plate method for studying root growth and response of Arabidopsis thaliana. *Sci Rep* 2013;3:1273. https://doi.org/10.1038/srep01273

Zabotina OA, Avci U, Cavalier D et al. Mutations in multiple XXT genes of Arabidopsis reveal the complexity of xyloglucan biosynthesis. *Plant Physiol* 2012;159:1367–84. https://doi.org/10.1104/pp.112.198119

Zur H, Ruppin E, Shlomi T. iMAT: an integrative metabolic analysis tool. *Bioinformatics* 2010;26:3140–2. https://doi.org/10.1093/bioinformatics/btq602