



Metabolic modelling revealed source–sink interactions between four segments of *Setaria viridis* leaves

RAKTIM MAITI¹, RAHUL SHAW^{2*}, C Y MAURICE CHEUNG³ and
SUDIP KUNDU^{1*}

¹Department of Biophysics, Molecular Biology and Bioinformatics, University of Calcutta, Kolkata, India

²Department of Plant Systems Physiology, Radboud University, Nijmegen, The Netherlands

³Yale-NUS College, 16 College Avenue West, Singapore, Singapore

*Corresponding authors (Emails, rahul.nutron@gmail.com; skbmbg@caluniv.ac.in)

MS received 17 September 2022; accepted 1 March 2023

As in most plants, during their growth from immature to mature stages, the leaves of *Setaria viridis*, a model C₄ bioenergy plant, have differential growth rates from the base (immature or growing) to the tip (most mature). In this study, we constructed a multi-segment C₄ leaf metabolic model of *S. viridis* with two cell types (bundle sheath and mesophyll cells) across four leaf segments (base to tip). We incorporated differential growth rates for each leaf segment as constraints and integrated transcriptomic data as the objective function for our model simulation using flux balance analysis. The model was able to predict the exchanges of metabolites between immature and mature segments of the leaf and the distribution of the activities of biomass synthesis across those segments. Our model demonstrated the use of a modelling approach in studying the source–sink relationship within an organ and provided insights into the metabolic interactions across different parts of a leaf.

Keywords. C₄ photosynthesis; flux balance analysis; genome-scale metabolic model; metabolic modelling; *Setaria viridis*

1. Introduction

In the process of photosynthesis, plants fix inorganic CO₂ from air to produce organic compounds. Among three major types of photosynthesis, C₄ photosynthesis involves two spatially separated cell types in the leaf called mesophyll (MP) and bundle sheath (BS) in tandem to concentrate CO₂ around RuBisCO in BS cells (Kanai and Edwards 1999), which leads to a reduction of RuBisCO oxygenation activity and photorespiration (Sage 1999), a process that can be considered wasteful. In C₄ plants, CO₂ first enters the MP and is fixed into a C₄ compound followed by the transport of a C₄ compound to the BS, where it is decarboxylated to provide a CO₂-rich environment around RuBisCO.

During the course of evolution, C₄ plants have gained several beneficial traits compared with their C₃ counterparts, such as high water- and nitrogen-use efficiencies and higher rates of productivity in hot and/or dry conditions (Sage 2004). *Setaria viridis*, a member of the Panicoideae clade, has an optimum growth temperature range of 20–35°C, relatively simple growth requirements, and a short life cycle (8–10 weeks). Additionally, *S. viridis* contains a comparatively small genome (~500 Mb) and large similarities with several major fuel and crop grasses, making it easy for genetic manipulations, thus establishing it as a model C₄ grass (Huang *et al.* 2016).

Plant leaves provide carbon and energy for growth and development of the whole plant, including the

Supplementary Information: The online version contains supplementary material available at <https://doi.org/10.1007/s12038-023-00348-0>.

leaves themselves. Leaf cells go through different developmental stages; thus, a single leaf can be divided into several segments of development. Each segment has a different metabolic behaviour and source–sink relation connecting the global leaf metabolic circuitry starting from the base, where cells proliferate, to the tip, where mostly cell expansion takes place (Kalve *et al.* 2014; Wu *et al.* 2021). Thus, although a leaf is often considered a single unit, in reality it has zonal biophysical divisions (Kalve *et al.* 2014; Wu *et al.* 2021) at different developmental stages (Zhang *et al.* 2021). This complex metabolic cross-talk at the level of demand and supply or source–sink relation within a C₄ leaf can be captured by using a systems biology approach. High-throughput sequencing of plant tissues can precisely identify diverse molecular heterogeneity within the same tissue along the developmental trajectory (NCBI-GEO:GSE89855, Zhang *et al.* 2021). Such transcriptomic data can be utilized to gain insights into the metabolic behaviour in different parts of an organ and help link differential gene expression to the leaf segments. Thus, our study can help evaluate the pseudo-time of gene targets (Cao *et al.* 2019) as well as real-time metabolic impact on leaf development.

Stoichiometric modelling of plant metabolism has been shown to be an efficient computational method to study systems-level metabolic properties (Poolman *et al.* 2009, 2013; Cheung *et al.* 2013). Previous computational studies on the metabolism of various C₄ plants including maize (Saha *et al.* 2011), a bioenergy plant *Setaria italica* (de Oliveira Dal’Molin *et al.* 2016), a model C₄ plant *Setaria viridis* (Shaw and Cheung 2019), and a general purpose C₄ model (de Oliveira Dal’Molin *et al.* 2010) have uncovered new metabolic properties. C₄ photosynthesis is a product of synergistic interactions between structural adjustments and a series of biochemical reactions and metabolite transport; i.e., it is not just a single metabolic pathway (Sage 2004). A computational study using a maize metabolic model successfully captured the relation of base-to-tip metabolic transition (Bogart and Myers 2016) by dividing the leaf into 15 segments. A high-resolution study using a model C₄ plant is lacking. Thus, in this study, based on a metabolic model of the *S. viridis* leaf (Shaw and Cheung 2019), we developed a multi-segment model capturing the developmental gradient from the growing to the differentiated states. We integrated segment-wise transcriptomic data with our multi-segment *Setaria* leaf model to explore the metabolic interactions between

leaf segments at different developmental stages and the gradient metabolic patterning across leaf segments. Similar to studies that use RNA sequencing of a leaf (Xie *et al.* 2020; Wang *et al.* 2021; Zhang *et al.* 2021) to depict a new regulatory pattern and role of genetic machinery in the order of development, our study helped dissect the role of leaf segments (a collection of cells in the same stage of development) while organizing C₄ leaf metabolic activity in mesophyll and bundle sheath cells. In this case, gene expression data (NCBI-GEO:GSE89855), which only highlight transcriptomic changes, when fit into a metabolic model also provided spatial resolution in the metabolic dissection of C₄ anatomy.

2. Materials and methods

2.1 Construction of multi-segment C₄ leaf metabolic model

We constructed a multi-segment C₄ leaf metabolic model of *S. viridis* based on a published genome-scale model (GSM) of *S. viridis* (Sv3376) (Shaw and Cheung 2019). For modelling C₄ photosynthesis, each leaf segment contained two types of cells, bundle sheath (BS) and mesophyll (MP). We duplicated the single-cell GSM into mesophyll and bundle-sheath cell types with the suffixes ‘MP’ and ‘BS’, respectively, similar to the previous *S. viridis* model (Shaw and Cheung 2019). All reactions and metabolites of the two-cell model were replicated four times to represent four different leaf segments, ‘base’, ‘lower’, ‘middle’, and ‘tip’ (using the suffixes ‘base’, ‘lwr’, ‘mid’, and ‘tip’, respectively), which correspond to the gradual transition from immature to mature portions of a typical *Setaria* leaf. The four leaf segments were connected to a common pool (‘CP’), representing xylem and phloem, which allowed exchange of metabolites between the leaf segments. We allowed exchange reactions for sucrose, 20 amino acids, and inorganic nutrients between the segments and the CP (Shaw and Cheung 2018). All these exchanges between the segments and CP were allowed through the BS (figure 1). This multi-segment metabolic model of a C₄ leaf (C4MS-GSM) of the model plant *S. viridis* (supplementary file 1) was validated for stoichiometric consistency, energy and redox conservation, and biomass production in all segments using our previous approaches (Shaw and Cheung 2018, 2019, 2021; Moreira *et al.* 2019).

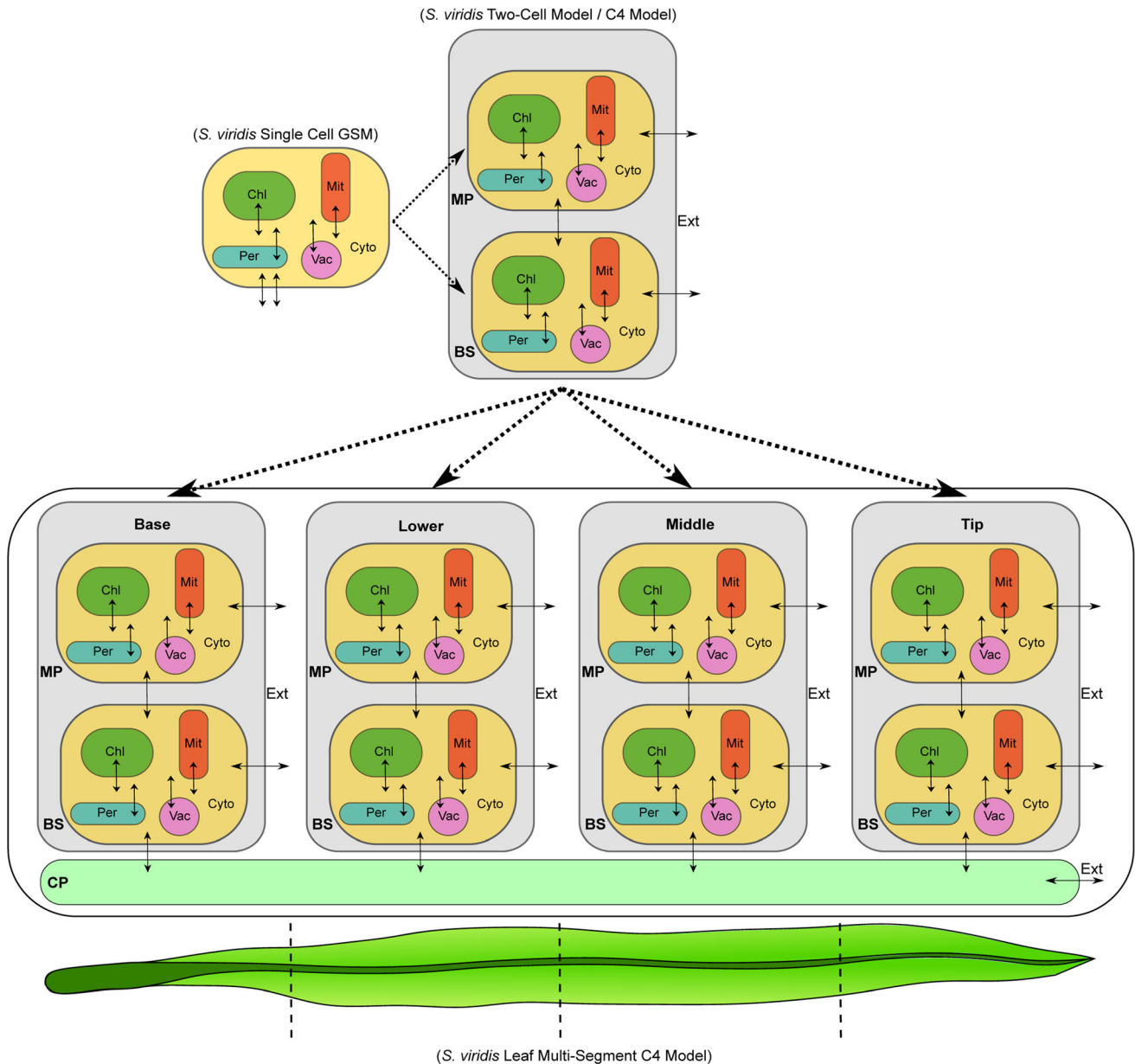


Figure 1. Reconstruction of multi-segment C₄ leaf model of *Setaria viridis* from the single-cell genome-scale model (GSM). The single-cell GSM was duplicated into a two-cell model representing bundle sheath (BS) and mesophyll (MP) cells for modelling C₄ photosynthesis. After adding transport reactions between the two cells, the two-cell model was further replicated four times, representing four leaf segments: Base, Lower, Middle, and Tip. The four segments were connected through a common pool (CP) representing the xylem and phloem. Model files are provided as supplementary file 1.

2.2 Model analysis using flux balance analysis (FBA) and flux variability analysis (FVA)

The *Setaria* C4MS-GSM model was analysed using flux balance analysis (FBA) (Orth *et al.* 2010). The aim of FBA was to derive a feasible set of steady state fluxes along with optimization of the objective function with given constraints. At steady state, the rate of production and the rate of consumption are equal for

each internal metabolite in the model, which can be represented mathematically as

$$S\mathbf{v} = 0$$

where \mathbf{v} is a vector of fluxes through the metabolic model and \mathbf{S} is the stoichiometry matrix of size $m \times n$. In the stoichiometric matrix, m represents the corresponding metabolites and n represents the corresponding reactions. We used FBA for validation and

analysis of our model before and after integration of transcriptomic data. For our FBA simulations, we used the scobra package (Shaw and Cheung 2018), which uses linear programming (LP) (using the Gnu Linear Programming Kit (GLPK), <http://www.gnu.org/software/glpk/>) to compute solutions of linear optimization problems. In our analysis, the linear optimization problem can be defined as: maximize or minimize $\mathbf{z} = \mathbf{w}^T \mathbf{v}$, subject to satisfying the steady state assumption and flux constraints

$$\begin{aligned} \mathbf{S}\mathbf{v} &= 0 \\ \mathbf{v}_l &\leq \mathbf{v} \leq \mathbf{v}_u \end{aligned}$$

Here, \mathbf{z} is the objective function, \mathbf{v} is the flux vector, \mathbf{w}^T is the transpose of vector of objective coefficients (\mathbf{w}) or weights representing the contribution of each reaction to the objective function, \mathbf{S} is the stoichiometry matrix, and \mathbf{v}_l and \mathbf{v}_u are the vectors of lower and upper bound of fluxes, respectively. While integrating transcriptomic data, we used the inverse of the gene expression value as the coefficient or weight (\mathbf{w}) of the corresponding reaction and minimization of the weighted sum of all reactions fluxes as the objective function (\mathbf{z}) (see section 2.4 for further details). We used a two-step optimization method to obtain a set of flux vectors from our model. In the first step, maximization of the biomass production was used as the objective function with a constrained photon influx of $200 \mu\text{mol m}^{-2}\text{s}^{-1}$ for each cell (Shameer et al. 2018) while all other input metabolites were unconstrained. In the second step, the flux for biomass production obtained from the first step was used as an additional constraint, and minimization of the absolute sum of all internal reaction fluxes weighted with corresponding coefficients was used as the objective to be optimized.

The *Setaria* C4MS-GSM model was further subjected to flux variability analysis (FVA). FVA checks allowable flux ranges of each reaction in the optimal flux solution space (Mahadevan and Schilling 2003; Thiele et al. 2010) with the given constraints and objective function. We used FVA to check the range of fluxes in our solution to make sure the conclusions made are not an artifact of alternative optimal solutions.

2.3 Biomass and other constraints

We utilized biomass data from the work of Bogart and Myers (2016) for maize given the unavailability of detailed stoichiometry of biomass composition for different leaf sections of *S. viridis*. Detailed

stoichiometric calculations of biomass components can be found in supplementary file 2.

Different segments of plant organs such as the root, stem, and leaf composed of zonal/segmental units have differential gene expression patterns, as well as diversity in regulatory and metabolic activities (Kalve et al. 2014; Li et al. 2016; Martin et al. 2016; Zhang et al. 2019; Sprangers et al. 2020). The meristem is composed of proliferating cells, whereas in the elongation segment, cells start to elongate and cell wall thickening occurs. In the maturation segment, cells grow to their final size with cell wall and cytoplasmic components. Based on the work of Bogart and Myers (2016), we divided the whole leaf into four segments by grouping three to four segments of Bogart and Myers (2016) into one to be consistent with the available transcriptomic data that we used. We compiled the biomass of those three to four segments into one to obtain segment-wise biomass components. Biomass equations in each cell type (MP and BS) in C4MS-GSM yield 1 mg of biomass $\text{m}^{-2}\text{s}^{-1}$ for one unit flux.

Each leaf segment was constrained to a segment-specific biomass synthesis rate as the base of the leaf produces more leaf biomass (i.e., growth) compared with the mature part of the leaf near the tip. The biomass production rates for the four leaf segments were calculated based on data from Bogart and Myers (2016). The biomass production constraints in the base, lower, middle, and tip segments were in the ratio of 5.4880:1.9372:1.2334:1.2830. Detailed calculations of the biomass constraints can be found in supplementary file 2. We set equal biomass production rates for bundle-sheath and mesophyll cells as the biomass production rates were found to be similar between the two cell types (Bogart and Myers 2016).

We also blocked the activities of chloroplastic NADPH dehydrogenase and plastoquinol oxidase given their insignificant contributions to photosynthesis (Wang and Fu 2016). Also, cytosolic NADP-ME (NADP-malic enzyme) and PEPCK (phosphoenolpyruvate carboxykinase) and NAD-ME (NAD-malic enzyme) in mitochondria were blocked in all segments with all other constraints being the same as in Shaw and Cheung (2019) for modelling the NADP-ME C₄ pathway.

2.4 Transcriptome data integration

To obtain more information on metabolism of leaf segments, we utilized segment-specific transcriptomic data of the *S. viridis* leaf from GEO dataset GSE89855 and integrated the transcriptomic data with C4MS-GSM.

The transcriptomic data were normalized using the TPM normalization method in the DESeq2 R-package (Love *et al.* 2014). Transcripts from the gene expression data were mapped to the *S. viridis* genome using reciprocal BLAST (Altschul *et al.* 1990; Camacho *et al.* 2009). The transcriptome data were used in the objective function of the flux balance model where the reciprocal of the gene expression values was used as the reaction weight in the objective function that was minimized. The expression of an enzyme has no linear relation with the corresponding reactions, but we followed the assumption that higher expression of a gene translates into higher production of the corresponding enzyme, which in turn allows a higher flux through the corresponding reaction, whereas lower expression of an enzymatic gene leads to lower availability of that enzyme, thus making the related reaction unfavourable. We used the mean of all gene expression values for the genes with unknown expression values. For reactions involving multiple genes, Gene-Protein-Reaction rules in the model were taken into considerations. If any of the multiple genes can facilitate a reaction, we took the gene with highest expression for our calculation; if multiple genes were essential for a reaction, the gene with lowest expression among the essential genes (rate-limiting enzyme) was taken for calculation of the weight. With this logic, a large gene expression value translates into a low weight of the reaction being minimized, i.e., favoured, compared with reactions with a higher weight (lower gene expression value) in the objective function which were penalized/unfavourable. To illustrate this concept, below is an illustrative example using arbitrary expression (E) values of five protein/enzyme coding genes; the weights (**w**) of the reactions in the objective function can be calculated, which indicate the preferential order of the five genes mapped to reactions (R):

Gene_A : E = 4.0 →

R_A(**w**_A = 1/E = 0.25) most favourable

Gene_B : E = 2.0 →

R_B(**w**_B = 1/E = 0.50) more favourable

Gene_C : E = 1.1 →

R_C(**w**_C = 1/E = 0.91) slightly favourable

Gene_D : E = 0.5 →

R_D(**w**_D = 1/E = 2.00) less favourable

Gene_E : E = 0.1 →

R_E(**w**_E = 1/E = 10.0) least favourable

The gene expression values and corresponding weights for each reactions in our model are given in

supplementary file 3. More details on our modelling methodologies can be found in the following github repository: <https://github.com/raktim2801/Svir-C4MS-GSM>.

3. Results

3.1 Model simulation of four leaf segments of *S. viridis*

The C4MS-GSM of *S. viridis* contains 19,687 reactions including intra- and inter-cellular transports and metabolite exchange reactions through a common pool (CP) representing the phloem and xylem. Each leaf segment contains 4,920 reactions inclusive of internal and external exchanges and transports. All reactions in the model, excluding external nutrients and biomass transport reactions are mass- and charge-balanced. We validated the energy conservation and stoichiometric consistency, so that biomass components cannot be produced without nutrients or energy. No ATP and NADPH production or electron transfer by transhydrogenation reactions could occur without external energy supply, making it a robust multi-segment leaf model to investigate C₄ leaf metabolism. A schematic representation of C4MS-GSM and model summary are given in figure 2 and table 1, respectively.

C4MS-GSM is capable of producing all biomass components (carbohydrates and lignin, amino acids, nucleic acids, and lipids) in different ratios utilizing photon energy, mineral nutrients (NH₄⁺, NO₃⁻, orthophosphate (Pi), SO₄²⁻, Mg²⁺), water, and protons. In the model, individual biomass reactions in each segment were composed of the unique stoichiometry of major biomass components. To represent realistic biomass proportions across segments, we ran our model simulations with the total biomass of the base, lower, middle, and tip segments fixed to a ratio of 5.4880:1.9372:1.2334:1.2830 (supplementary file 3). The photon flux was constrained to 200 μmol m⁻²s⁻¹ for each cell (Shameer *et al.* 2018) during the simulation (figure 2). For simulation of a C₄ leaf-specific metabolism, exchange of external CO₂ and O₂ were only allowed through the mesophyll along with exchange of selected carbon compounds and O₂ between mesophyll and bundle-sheath cells. Before the integration of gene expression data, we simulated our initial model with an objective function of maximizing biomass production in all segments constrained with a fixed photon input in the first optimization step, followed by minimization of the absolute sum of all

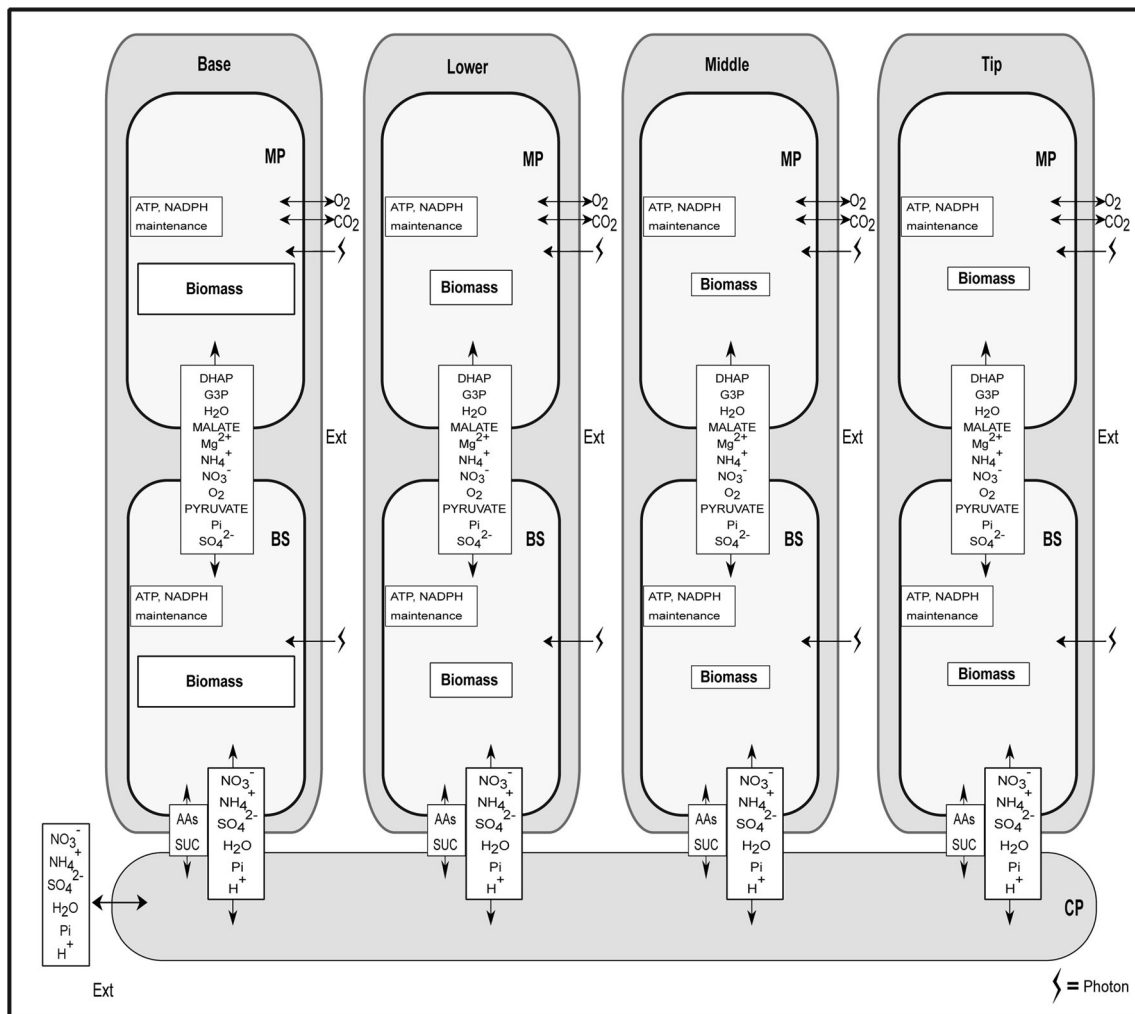


Figure 2. A schematic representation of the model indicating biomass and maintenance constraints and inter-cellular metabolite transports and exchanges between bundle sheath (BS) cells, mesophyll (MP) cells, and the common pool (CP). The area of ‘Biomass’ in the figure is scaled to the biomass synthesis rate of the respective cell.

reaction fluxes as the second step objective with the additional constraint of the biomass fluxes obtained in the first step. We observed that external CO_2 is assimilated into oxaloacetic acid (OAA) as the first organic compound in this model as part of the C_4 photosynthetic pathway. The model also utilized NADP-ME pathway for carbon assimilation in bundle-sheath cells, which is consistent with the previously known behaviour of *Setaria* (Brutnell et al. 2010), implying that the multi-segment model is a valid model for modelling C_4 photosynthesis in *S. viridis*. Flux distributions for all reactions of this model with and without gene expression integration are given in supplementary file 4. A representation of the flux distribution pattern before integration of transcriptomic data through the C_4 pathway in all segments can be found in supplementary file 5. We also obtained flux ranges from flux variability analysis (FVA) for this model

testing step, which confirms the robustness of the model prediction, e.g., the operation of C_4 photosynthesis in all segments. FVA solutions are given in supplementary file 6.

3.2 Integration of gene expression data gave realistic metabolite transport patterns between leaf segments

Towards simulating realistic differential metabolism across leaf segments, we incorporated gene expression (GE) data of *S. viridis* leaf (GEO dataset GSE89855) into the model by using the gene expression data in the objective function (see Methods for details; flux distribution is given in supplementary file 4). The integration of GE data resulted in several changes in the overall flux distribution pattern, which are reflected by

Table 1. Summary of the multi-segment *Setaria viridis* C₄ leaf mode (model files are provided as supplementary file 1)

Reactions	No. of reactions	Metabolites	No. of metabolites
Total	19687	Total	19490
Base	4920	Base	4866
Lower	4920	Lower	4866
Middle	4920	Middle	4866
Tip	4920	Tip	4866
External nutrients through CP	7	CP	26
MP	9728	Genes	No. of genes
BS	9728	—	—
MP-BS transporter	48	Unique	3376
CP-BS transporter	152		
External-CP transporter	7		
Cytoplasmic	14780		
Mitochondrial	656		
Plastidial	2080		
Peroxisomal	248		
External nutrient transporter	67		
Biomass equations	40		
CP transporter	152		
Internal transporter:			
mc	592		
vc	208		
pc	568		
xc	256		

the exchanges of metabolites through the CP. In the GE-integrated simulation, the main source of sucrose in CP changes from the middle segment to the tip (figure 3A). By incorporating the GE data, we can observe that the export of sucrose gradually increased from the lower segment to the tip, consistent with the maturity level of the leaf tissue, giving us a more biologically relevant flux distribution.

The flux distribution predicted in each segment showed the production of sucrose through the NADP-ME subtype of the C₄ photosynthetic pathway in all four leaf segments (figure 3B), consistent with the evidence in the literature (Brutnell *et al.* 2010). We also observe that the biomass demand of the base could not be fulfilled by its own carbon fixation, and was supplemented by the import of sucrose from the other three segments, as discussed below (figure 4A and B).

3.3 Metabolic interactions between leaf segments

The flux distribution of the CP exchange reactions demonstrated that sucrose was the major component of inter-segment exchanges (figure 4B and supplementary file 7). The base segment, being the segment with the greatest amount of growth, is expected to have high carbon and energy demands, which was reflected in the

flux distribution pattern. Along with sucrose, several other amino acids were imported into the base (figure 4B), where phenylalanine and glutamate were predicted to be the most abundant. The differences between the biomass needed for growth (constrained biomass) and the metabolic fluxes through biomass synthesis reactions (synthesized biomass) for each segment showed that the base segment fulfils a significant portion of its biomass demand by importing biomass components from other parts of the leaf (figure 4A). We also observed that the amount of sucrose, phenylalanine, and glutamine imported into the base is higher than the amount needed for its biomass (supplementary file 7), indicating that those metabolites were utilized for the production of other biomass components and/or as an energy source in the base segment.

3.4 Metabolic differences between four segments of C₄ leaf

A comparison between the active reactions predicted among the four segments is represented as a Venn diagram in supplementary file 8. A large percentage of active reactions are common in all the segments (565 reactions, 74.9%). Reactions common to the lower,

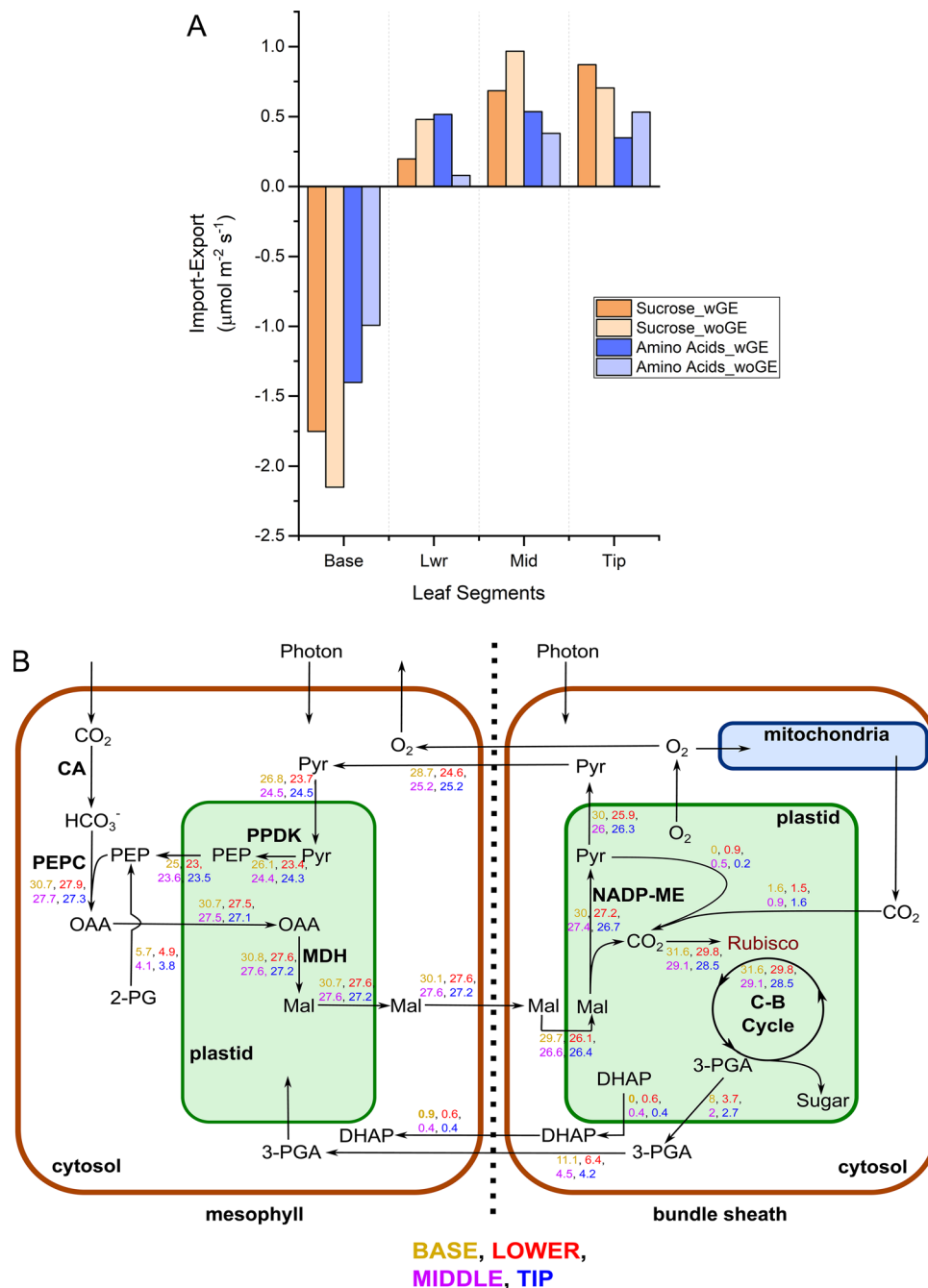


Figure 3. (A) Comparison of metabolite exchanges between simulations with integration of gene expression data (wGE) and without gene expression data (woGE). The bar chart represents net import or export of sucrose and amino acids into or out of each leaf segment with positive values representing exports and negative values representing imports. (B) Flux distribution across mesophyll and bundle-sheath cells, indicating carbon fixation through the NADP-ME subtype of C₄ photosynthesis in all four leaf segments modelled. Flux values of the base, lower, middle, and tip segments are shown in yellow, red, purple, and blue, respectively. Distribution of fluxes follows the same qualitative pattern in each segment. Flux distributions are given in supplementary file 4.

middle, and tip segments formed the next largest set (89 reactions, 11.8%), suggesting that there is more in common between those three segments, with the base segment being the most different. This was followed by reactions unique to the lower (35 reactions, 4.6%) and

base (24 reactions, 3.2%) segments. Looking further into two mutually exclusive sets, unique reactions in the base and common reactions in the lower, middle, and tip segments, we observed that the main differences between them are related to the import and

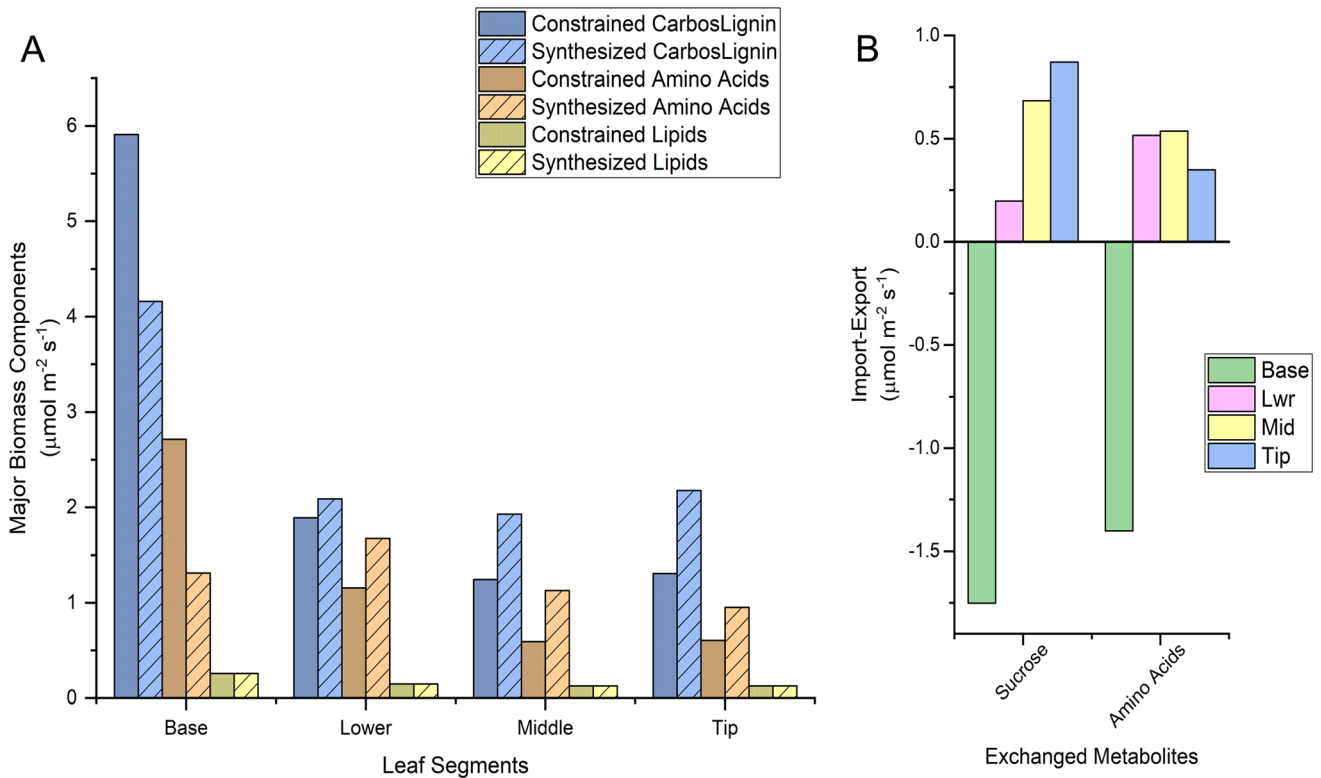


Figure 4. (A) Bar chart representing major biomass components (y-axis) produced at different segments of the leaf (x-axis). Here, we differentiated biomass synthesis into two categories: (i) Constrained, representing the amount of biomass for growth (i.e., values set for the biomass constraints of the model) and (ii) Synthesized, representing the amount of biomass synthesised *de novo* in each segment calculated as Constrained biomass – Imported biomass + Exported biomass. CarboLignin, Amino acids, and Lipids are the major biomass components shown in this diagram: (i) CarboLignin = carbohydrates + lignin, (ii) Amino Acids = free amino acids + proteins, and (iii) Lipids = lipids + fatty acids. (B) Comparison of exchanges of metabolites between leaf segments through the common pool (CP). Exports of metabolites from the leaf segments to the CP are represented as positive values, whereas imports of metabolites from the CP into the leaf segments are represented as negative values. Base (green), Lwr (pink), Mid (yellow), and Tip (blue) represent the base, lower section, middle section, and the tip of the leaf, respectively.

export of sucrose. Bundle-sheath cells of the lower, middle, and tip segments can synthesize their carbohydrate biomass components by themselves and also fulfil the carbon demands of the corresponding mesophyll cells by sharing 3-C compounds across the plasmodesmata (figure 5). The excess carbohydrates get exported to the CP from these three segments, whereas in the base, photosynthesis alone is not sufficient to fulfil its carbon demand. It imports sucrose from other segments via the CP for producing part of its biomass components. Additionally, most of the unique reactions in the lower segment are associated with the synthesis of a few amino acids (isoleucine, methionine, threonine) (supplementary file 9). The flux distribution indicates that the demand for these amino acids in other segments is also fulfilled by the lower segment.

4. Discussion

4.1 Comparison with multi-section models

Photosynthesis, a complex biological process, involves high-magnitude activity of the enzymatic machinery for energy production, energy use, and CO₂ assimilation, where the C₄ type has additional steps involving spatial separation. Constraint-based models in the past decades have highlighted static and dynamic roles of different interactions of the photosynthetic metabolism with a holistic view (de Oliveira Dal’Molin *et al.* 2010; Poolman *et al.* 2013; Chatterjee and Kundu 2015; Chatterjee *et al.* 2017; Huma *et al.* 2018). Most of the models were focused on the C₃ single-leaf cell or generalized leaf model which were sufficient to explore metabolic dynamics as a whole. However, studies on

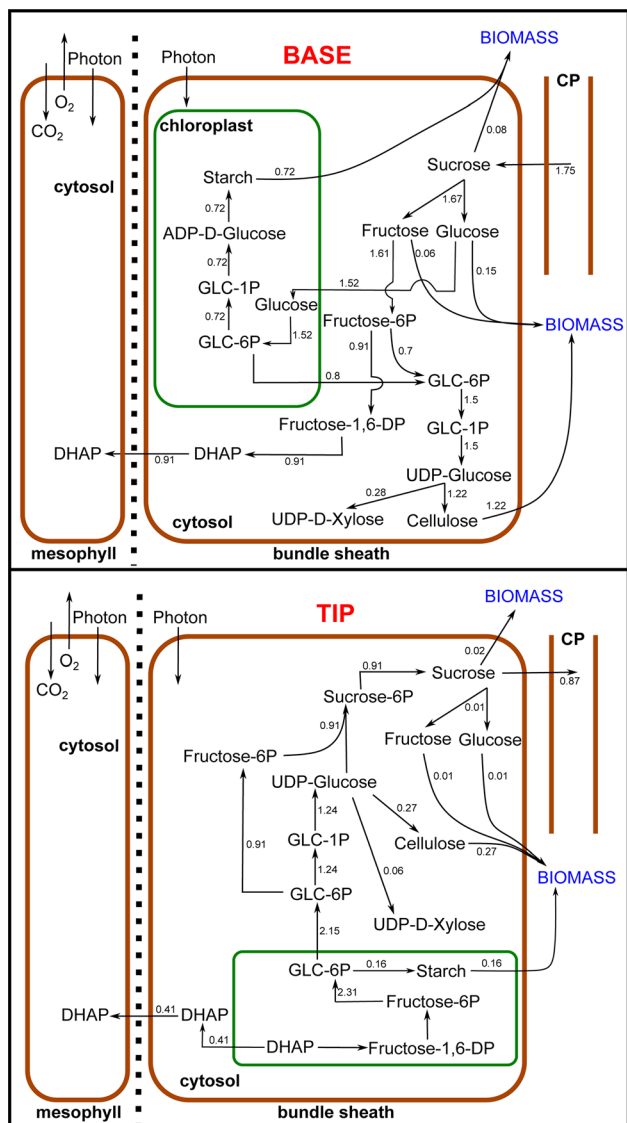


Figure 5. Comparison of predicted flux distribution patterns between leaf segments. Flux values of only the base and tip segments are shown in the figure. The lower and middle segments have a flux pattern qualitatively identical to the tip segment. Note that the base imports sucrose from the common pool while the other three leaf segments export sucrose to the common pool. Flux distributions are given in supplementary file 4.

the photosynthesis within the developmental aspects of a single leaf are scarce (Bogart and Myers 2016). To the best of our knowledge, only Bogart and Myers (2016) introduced differentiating units of a leaf and their interactions within a model. Introduction of the organ differentiation dimension for a C_4 photosynthetic leaf using the model organism *S. viridis* in our present study explores the complex interactions for resource and energy within a leaf. This model is also able to predict the exchanges of metabolites between the

immature and mature leaf segments, and predicts the biomass synthesis across those segments. This will provide a model-based approach to understand the source–sink relationship within a plant organ and the metabolic interactions between different leaf segments.

4.2 Modelling framework for modelling metabolic interactions between leaf segments

A typical plant leaf consists of many cells at different growing stages starting from proliferating cells near the base to the differentiated cells at the tip. Each developmental stage demands its own metabolic phenotype. High-throughput sequencing studies such as single-cell RNA sequencing highlight a robust heterogeneity spread over a single plant organ (Ryu *et al.* 2019; Zhang *et al.* 2019, 2021). A growing plant organ, such as a root, has developmental zone- or even fine section-level dominant expression patterns along its longitudinal axis (Brady *et al.* 2007). Similarly, along the longitudinal segments, a leaf also consists of cells at different developmental stages (Kalve *et al.* 2014). These tissue segments are also bound to have high to marginal levels of metabolic variability along with differential gene expression patterns (Tausta *et al.* 2014). Similar to different tissues with different metabolic needs and activities, we can expect there to be metabolic source–sink relationships among these segments (Ludewig and Flüge 2013; Griffiths *et al.* 2016). In our model, we used differential biomass along with segment-specific expression of enzymatic genes to simulate the whole leaf at a single point of time, where different cell types are separated spatially rather than temporally. This simulation demonstrates a more realistic physiological scenario of C_4 plant leaf metabolism, where cells at different developmental stages fulfil their unique metabolic demands by exchanging necessary metabolites among them at any given time, without affecting the C_4 pathway of carbon assimilation. This is a step towards more complex and realistic models and simulations of growth patterns with integration of differential external conditions. Our framework can help establish relationship of gene target(s) to specified growth time-points, such as pseudo-time of cell progression to a specific differentiation limit and real-time of leaf development stages. In this study we also used a novel approach for integration of transcriptomic data, based on simple assumptions as mentioned in the Methods section. The expression of a gene does not linearly correlate with the flux through a reaction, but the gene expression indirectly influences

the feasibility of enzymes to catalyse reactions to a certain degree. Hence, we used the expression data as an influencing factor in our approach instead of establishing direct correlations with the reaction flux values.

4.3 Division of labour of biomass synthesis between leaf segments

Our model predicted that sucrose was the major component transported between the leaf segments, consistent with the phloem sap metabolic profile of the model plant *Arabidopsis* (Tetyuk *et al.* 2013). Interestingly, our model predicted the use of phenylalanine as one of the major components being transported between leaf segments. Phenylalanine is observed to increase in phloem sap of water-deficit treated *Eucalyptus globulus*, a C₃ plant (Merchant *et al.* 2010), which may indicate that the predicted use of phenylalanine in the phloem of C₄ *Setaria* plants could be related to water use efficiency. Results from flux variability analysis (FVA) showed that sucrose transport to the phloem and utilization of phenylalanine across leaf segments had no flux variability (supplementary file 6), which means the FBA prediction was robust and not subjected to alternative flux solutions. The model also suggests that the lower, mid, and tip segments need to supply sucrose and amino acids to the base to meet its high growth and energy demand (figure 4A). In a plant body, mature leaves are the main source of photosynthates for the whole plant, but a single leaf can also show such a source–sink relationship with the mobilization of sugar and other biomass precursors to fulfil local or metabolically transient demands. Gene expression-integrated simulation highlights the specialized roles of the lower and mid segments to provide more amino acids to the base (figures 3A and 4A) than the tip. This could reflect a spatial relationship between the neighbouring segments for fulfilling faster supply of protein building blocks without imposing the burden of additional transport cost and/or distance from the tip. However, the preference for sucrose supply from the tip instead of the middle segment under GE-integrated simulation (figure 3A) might indicate that sugar export directly from the most mature segment is preferred. Our model showed that nearly half of the demand of protein synthesis in the base is fulfilled by import (figure 4A) whereas lipid demand can meet by local biosynthesis (figure 4A), which is consistent with the observation that fatty acids are synthesized locally in maize immature leaf tissue (Hawke *et al.* 1974). These

results were confirmed using FVA, which showed no or very small flux variability (small flux ranges) for the processes described above (supplementary file 6), suggesting that the results were not an artefact of alternative optimal solutions.

Acknowledgements

RS thanks the Radboud University for his fellowship.

Author contributions

CYMC formulated the study design. RM reconstructed the model and performed the simulations. RM, RS, SK and CYMC analysed the results and co-wrote the manuscript.

Funding

Yale-NUS College.

Declarations

Conflict of interest The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Altschul SF, Gish W, Miller W, *et al.* 1990 Basic local alignment search tool. *J. Mol. Biol.* **215** 403–410
- Bogart E and Myers CR 2016 Multiscale metabolic modeling of C₄ plants: connecting nonlinear genome-scale models to leaf-scale metabolism in developing maize leaves. *PLoS One* **11** e0151722–e0151722
- Brady SM, Orlando DA, Lee J-Y, *et al.* 2007 A high-resolution root spatiotemporal map reveals dominant expression patterns. *Science* **318** 801–806
- Brutnell TP, Wang L, Swartwood K, *et al.* 2010 *Setaria viridis*: a model for C₄ photosynthesis. *Plant Cell* **22** 2537–2544
- Camacho C, Coulouris G, Avagyan V, *et al.* 2009 BLAST+: architecture and applications. *BMC Bioinform.* **10** 421
- Cao J, Spielmann M, Qiu X, *et al.* 2019 The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566** 496–502
- Chatterjee A and Kundu S 2015 Revisiting the chlorophyll biosynthesis pathway using genome scale metabolic model of *Oryza sativa japonica*. *Sci. Rep.* **5** 14975
- Chatterjee A, Huma B, Shaw R, *et al.* 2017 Reconstruction of *Oryza sativa indica* genome scale metabolic model and

- its responses to varying RuBisCO activity, light intensity, and enzymatic cost conditions. *Front. Plant Sci.* **8** 2060–2060
- Cheung CYM, Williams TCR, Poolman MG, et al. 2013 A method for accounting for maintenance costs in flux balance analysis improves the prediction of plant cell metabolic phenotypes under stress conditions. *Plant J.* **75** 1050–1061
- Dale JE 1992 How do leaves grow? *Bioscience* **42** 423–432
- de Oliveira Dal'Molin CG, Quek L-E, Palfreyman RW, et al. 2010 C4GEM, a genome-scale metabolic model to study C₄ plant metabolism. *Plant Physiol.* **154** 1871–1885
- de Oliveira Dal'Molin CG, Orellana C, Gebbie L, et al. 2016 Metabolic reconstruction of *Setaria italica*: a systems biology approach for integrating tissue-specific omics and pathway analysis of bioenergy grasses. *Front. Plant Sci.* **7** 1138–1138
- Grafahrend-Belau E, Junker A, Eschenröder A, et al. 2013 Multiscale metabolic modeling: dynamic flux balance analysis on a whole-plant scale. *Plant Physiol.* **163** 637–647
- Griffiths CA, Paul MJ and Foyer CH 2016 Metabolite transport and associated sugar signalling systems underpinning source/sink interactions. *Biochim. Biophys. Acta* **1857** 1715–1725
- Hawke JC, Rumsby MG and Leech RM 1974 Lipid biosynthesis in green leaves of developing maize. *Plant Physiol.* **53** 555–561
- Huang P, Shyu C, Coelho CP, et al. 2016 *Setaria viridis* as a model system to advance millet genetics and genomics. *Front. Plant Sci.* **7** 1781–1781
- Huma B, Kundu S, Poolman MG, et al. 2018 Stoichiometric analysis of the energetics and metabolic impact of photorespiration in C₃ plants. *Plant J.* **96** 1228–1241
- Kalve S, De Vos D and Beemster GTS 2014 Leaf development: a cellular perspective. *Front. Plant Sci.* **5** 362–362
- Kanai R and Edwards GE 1999 The biochemistry of C₄ photosynthesis; in *C₄ plant biology* (Eds) RF Sage and RK Monson (Academic Press) pp. 49–87
- Li S, Yamada M, Han X, et al. 2016 High-resolution expression map of the *Arabidopsis* root reveals alternative splicing and lincRNA regulation. *Dev. Cell.* **39** 508–522
- Love MI, Huber W and Anders S 2014 Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15** 550
- Ludwig F and Flügge U-I 2013 Role of metabolite transporters in source-sink carbon allocation. *Front. Plant Sci.* **4** 231–231
- Mahadevan R and Schilling CH 2003 The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metab. Eng.* **5** 264–276
- Martin AP, Palmer WM, Brown C, et al. 2016 A developing *Setaria viridis* internode: an experimental system for the study of biomass generation in a C₄ model species. *Biotechnol. Biofuels* **9** 45
- Merchant A, Peuke AD, Keitel C, et al. 2010 Phloem sap and leaf delta13C, carbohydrates, and amino acid concentrations in *Eucalyptus globulus* change systematically according to flooding and water deficit treatment. *J. Exp. Bot.* **61** 1785–1793
- Moreira TB, Shaw R, Luo X, et al. 2019 A genome-scale metabolic model of soybean (*Glycine max*) highlights metabolic fluxes in seedlings. *Plant Physiol.* **180** 1912–1929
- Orth JD, Thiele I and Palsson BO 2010 What is flux balance analysis? *Nat. Biotechnol.* **28** 245–248
- Poolman MG, Miguet L, Sweetlove LJ, et al. 2009 A genome-scale metabolic model of *Arabidopsis* and some of its properties. *Plant Physiol.* **151** 1570–1581
- Poolman MG, Kundu S, Shaw R, et al. 2013 Responses to light intensity in a genome-scale model of rice metabolism. *Plant Physiol.* **162** 1060–1072
- Ryu KH, Huang L, Kang HM, et al. 2019 Single-cell RNA sequencing resolves molecular relationships among individual plant cells. *Plant Physiol.* **179** 1444–1456
- Sage RF 1999 Why C₄ photosynthesis? in *C₄ Plant Biology* (Eds) RF Sage and RK Monson (Academic Press) pp. 3–16
- Sage RF 2004 The evolution of C₄ photosynthesis. *New Phytol.* **161** 341–370
- Saha R, Suthers PF and Maranas CD 2011 *Zea mays* iRS1563: a comprehensive genome-scale metabolic reconstruction of maize metabolism. *PLoS One* **6** e21784–e21784
- Shameer S, Baghalian K, Cheung CYM, et al. 2018 Computational analysis of the productivity potential of CAM. *Nat. Plants* **4** 165–171
- Shaw R and Cheung CY 2018 A dynamic multi-tissue flux balance model captures carbon and nitrogen metabolism and optimal resource partitioning during *Arabidopsis* growth. *Front. Plant Sci.* **9** 884
- Shaw R and Cheung CYM 2019 A mass and charge balanced metabolic model of *Setaria viridis* revealed mechanisms of proton balancing in C₄ plants. *BMC Bioinform.* **20** 357
- Shaw R and Cheung CY 2021 Integration of crop growth model and constraint-based metabolic model predicts metabolic changes over rice plant development under water-limited stress. *in silico Plants* **3** diab020
- Simons M, Saha R, Amiour N, et al. 2014 Assessing the metabolic impact of nitrogen availability using a compartmentalized maize leaf genome-scale model. *Plant Physiol.* **166** 1659–1674
- Sprangers K, Thys S, van Dusschoten D, et al. 2020 Gibberellin enhances the anisotropy of cell expansion in the growth zone of the maize leaf. *Front. Plant Sci.* **11** 1163–1163
- Tausta SL, Li P, Si Y, et al. 2014 Developmental dynamics of Kranz cell transcriptional specificity in maize leaf reveals early onset of C₄-related processes. *J. Exp. Bot.* **65** 3543–3555

- Tetyuk O, Benning UF and Hoffmann-Benning S 2013 Collection and analysis of *Arabidopsis* phloem exudates using the EDTA-facilitated method. *J. Vis. Exp.* **80** e51111
- Thiele I, Fleming RMT, Bordbar A, *et al.* 2010 Functional characterization of alternate optimal solutions of *Escherichia coli*'s transcriptional and translational machinery. *Biophys. J.* **98** 2072–2081
- Wang D and Fu A 2016 The plastid terminal oxidase is a key factor balancing the redox state of thylakoid membrane. *Enzymes* **40** 143–171
- Wang Y, Huan Q, Li K, *et al.* 2021 Single-cell transcriptome atlas of the leaf and root of rice seedlings. *J. Genet. Genomics* **48** 881–898
- Wu W, Du K, Kang X, *et al.* 2021 The diverse roles of cytokinins in regulating leaf development. *Hortic. Res.* **8** 1–13
- Xie Y, Jiang S, Li L, *et al.* 2020 Single-cell RNA sequencing efficiently predicts transcription factor targets in plants. *Front. Plant Sci.* **11** 1946–1946
- Zhang T-Q, Xu Z-G, Shang G-D, *et al.* 2019 A single-cell RNA sequencing profiles the developmental landscape of *Arabidopsis* root. *Mol. Plant* **12** 648–660
- Zhang T-Q, Chen Y and Wang J-W 2021 A single-cell analysis of the *Arabidopsis* vegetative shoot apex. *Dev. Cell* **56** 1056–1074

Corresponding editor: UTPAL NATH