Millet-inspired systems metabolic engineering of NUE in crops

Piyoosh K. Babele,[1,*] Amit Srivastava,[2,3] Khaled A. Selim,[4] and Anil Kumar[1]

[1]Rani Lakshmi Bai Central Agricultural University, Jhansi 284003, Uttar Pradesh, India

[2]Department of Chemistry, Purdue University, West Lafayette, Indiana 47907, United States of America

[3,#]University of Jyväskylä, Nanoscience Centre, Department of Biological and Environmental Science, 40014 Jyväskylä, Finland.

[4]Interfaculty Institute for Microbiology and Infection Medicine, Organismic Interactions Department, Tübingen University, Germany

*****Correspondence*:* piyoosh.bhu06@gmail.com (P.K. Babele)

#Current affiliation

***Keywords*:** nitrogen use efficiency; millets; flux balance analysis; metabolic flux analysis; PII protein.

Abstract

The use of nitrogen fertilizers in agriculture has a great ability to increase crop productivity. However, their excessive use has detrimental effects on the environment. Therefore, it is necessary to develop crop varieties with improved nitrogen use efficiency (NUE) that require less nitrogen but have substantial yields. Orphan crops such as millets are cultivated in limited regions and are well adapted to lower input conditions. Therefore, they serve as a rich source of beneficial traits that can be transferred into major crops to improve their NUE. This review highlights the tremendous potential of systems biology to unravel the enzymes and pathways involved in the nitrogen metabolism of millets, which can open new possibilities to generate transgenic crops with improved NUE.

Nitrogen use efficiency in millets

The use of nitrogen (N) fertilizers is a common practice in agriculture to improve crop productivity, but due to their excess application, they have become the largest anthropogenic source of reactive N in the environment (100 Tg per year)[1]. Reactive N has been identified as one of the major emerging environmental threats with cascading effects on ecosystems, biodiversity, and climate (UNEP, 2019)[2]. Currently, there is a pressing need to find alternatives and sustainable solutions to minimize the use of N in the soil while still maintaining substantial crop productivity. Designing genetically engineered crops with high **NUE** (see Glossary) could be a sustainable alternative to the overuse of N in agriculture. In order to generate genetically modified **major crops** (i.e., staple crops), which are usually a more popular source of food worldwide, it will be important to first investigate and characterize the molecular basis of NUE in those crops that flourish well under lower N conditions. A wide range of such crops with efficient N utilization belong to a group commonly known as minor or **orphan crops**. These crops are only regionally important and are not traded globally, but due to their adaptation to the local environment, they have evolved distinct adaptation strategies to overcome nutrient limitations [6]. **Systems biology**, which offers a comprehensive depiction of the intricate interactions that occur inside cellular metabolism, is a promising method to achieve a deeper insight into the mechanism by which orphan crops utilize N effectively. These analyses are carried out by an interdisciplinary method using computational, mathematical, and multi-omics tools. By monitoring the metabolic flow of cellular flux, it is possible to infer which enzymes or pathways (and their genes) are activated in a given environment (e.g., N limitation). If the candidate genes identified using such methods, are then transferred to major crops *via* synthetic biology tools, **metabolically engineered cultivars** with higher NUE can be generated. This combination of systems biology, synthetic biology, and metabolic engineering is termed "**systems metabolic engineering**," (Figure 1) which is gaining significant popularity in the engineering of microbial cell factories but has remained underutilized in agricultural biotechnology [3].

Many orphan crops, usually represented by millets, are **C4 plants** (Figure 2) and are tolerant to adverse conditions such as high temperature, low moisture, and limited nutrients [4, 5]. It has been long established that C4 plants are capable of utilizing N more efficiently than **C3 plants** (e.g., wheat, rice maize) (Figure 2) [5, 6]. In addition, C4 plants also use water- and light more efficiently when compared to C3 plants [7]. The higher NUE of C4 plants has been attributed to more efficient N distribution within the plant as well as possible spatial separation of **photorespiration** and nitrate assimilation reactions. Photosynthetic carboxylation enzymes such as RuBisCO (ribulose-1,5-bisphosphate carboxylase/oxygenase) are a major N-storing enzyme in plant cells [8]. Since, C4 species need less

RuBisCO than C3 plants due to their **CO_2 concentration mechanism**, they typically invest 3-4 times less N in RuBisCO than C3 plants (5-10% *vs*. 20-30%) [9]. Consequently, RuBisCO in C4 plants accounts for only 20% of the stored N compared to 50% in C3 plants. As N-storing RuBiCo activity is crucial to the ability of plants to perform photosynthesis, it is conceivable that C3 plants will be more negatively impacted by a decrease in N availability than C4 plants [9]. Therefore, it is possible that the ability of millets to utilize N much more efficiently is mainly attributed to their C3 metabolism. While it is established that millet display higher NUE than C3 crops, the details on the genetic and metabolic components behind these traits are largely unknown and require the implementation of modern systems biology methods to obtain a complete depiction of their NUE mechanism [10, 11]. In this review, we discuss the NUE associated characteristics of millets and describe the methods in systems biology to pinpoint their key genes, enzymes, and pathways of N metabolism. This review also covers several aspects of metabolic engineering that can be considered for delivering and engineering millet genes into major crops.

Systems biology of nitrogen metabolism in millets

The traditional method of genetic manipulation, which depends on database curation and gene function prediction (based on sequence similarity), frequently fails to provide the desired phenotypes or meaningful data. This is because the traditional methods are dependent on the so-called reductionist approach, which focuses on a single gene rather than taking into account the intricate and dynamic biological interactions these genes (and their enzymes) have with other biomolecules in the cell. Further, as the total metabolic activity of organisms is a result of the combined action of numerous genes, chemical equilibriums, and multiple layers of regulation (transcriptional, post-transcriptional, and post-translational), they add further complexity to their metabolic systems [12]. To study the complex interactions between NUE and other plant metabolisms, modern analytical approaches and techniques to capture a detailed picture of steady-state metabolite concentrations and associated **metabolic fluxes** are required. Therefore, to effectively transfer the NUE-associated metabolic potential of millets into staple crops, the first step would be to unravel the entire metabolic network of enzymes, regulators, and pathways in millets. This picture of complete metabolic activity could help to identify various interactions among pathways that could define the rate-limiting steps, bottlenecks, and feedback regulations among the metabolic pathways of an organism. Recent research provides encouraging evidence that it is feasible to create functional transgenic C3 plants using genes from C4 plants. For instance, it was found that overexpressing the SiMYB19 transcription factor from foxtail millet (C4 plant) promotes the accumulation of abscisic acid in transgenic rice (C3 plants) and increases the

expression of the abscisic acid (ABA) synthesis gene OsNCED3 as well as genes related to the ABA signal transduction pathway, OsPK1 and OsABF2. SiMYB19 controls ABA production and signal transduction, hence enhancing salt tolerance in transgenic rice [13].

Constructing **genome-scale metabolic models** (GSM) is one of the first steps in successfully implementing systems biology approaches in an organism [14]. The first step in the curation of a plant GSM is to identify all the genes present in the target crop species (Table 1). Usually, this process starts with the annotation of a genome and the prediction of candidate metabolic functions at a genome-scale. These GSMs can be used to simulate metabolic fluxes for various systems-level metabolic research and to computationally define gene-protein-reaction relationships for all of the metabolic genes of an organism. The process of GSM reconstruction and validation involves creating an initial model, curating the model, and then rendering the species- and context-specific models ready for **flux balance analysis** (FBA) [15]. The FBA, which comprehensively predicts cellular metabolic flux distributions for a given genotype and given environmental conditions, enables the quantitative visualization of carbon (C) flow through metabolic pathways. These flux predictions facilitate the hypothesis of new network properties. Eventually, **metabolic flux analysis** (MFA) is performed, which requires estimation of metabolites, which is done by nuclear magnetic resonance (NMR) and mass spectroscopy (MS). Modern MFA of plants relies on two possible approaches: steady-state MFA (^{13}C-MFA) and isotopically nonstationary metabolic flux analysis (INST-MFA), where metabolites are labeled with stable isotopes (^{13}C, ^{15}N). In particular, the **metabolomics** and **fluxomics** of photoautotrophic species have been transformed by INST-MFA [16, 17]. The photoautotrophs acquire C from CO_2 and create a homogeneous steady-state ^{13}C-labeling pattern that becomes insensitive to fluxes, making it difficult to trace the movement of radiolabeled C in the metabolic pathway. To overcome this problem, INST-MFA can be used for temporary detection of isotope incorporation after the switch from CO_2 to $^{13}CO_2$ supplementation. So far, this technique has been applied to model cyanobacteria and plants, but in the near future, INST-MFA could be a key player in the exploration of N metabolism of millets [16, 18]. Altogether, these techniques usually quantify fluxes based on fitting experimental metabolite data, consisting of external rates and isotope labeling patterns, to a core metabolic network model. Many software tools (INCA, Metran) are available to perform MFA calculations and statistical analysis by fitting the estimated flux with 95% confidence intervals [19]. The outcome of GSM and FBA analysis can then be used to estimate the fluxes through large metabolic pathways and then generate a simplified model for NUE. MFA quantifies the flow of metabolites through a core metabolic pathway, yielding flux maps that can aid engineering efforts by explaining phenotypes in detail. Recently, a mass and charge balance GSM of

Foxtail millet (*Setaria viridis*) was constructed and tested to be able to produce major biomass components [20]. This study demonstrated the use of metabolic modeling in identifying genes associated with the synthesis of particular biomass components and provides the possibility of system-level investigation to identify metabolic characteristics based on stoichiometric constraints. The manual reconstruction procedure often used to create a GSM is typically labor-intensive. However, there are various automated methods for genome annotation accessible, including ModelSEED, RAVEN 2.0 and Merlin that help to simplify this process [14]. The user then iteratively refines the initial reconstruction by thoroughly examining each reaction, metabolite, and gene in the network. Recent developments in **constraint-based reconstruction and analysis** (COBRA) have shown the potential to reveal the guiding principles of C4 carbon and N metabolic pathways and their interactions with other parts of the metabolism on a system-level (Table 2) [15]. In COBRA, data-driven physicochemical and biological constraints are used to enumerate the range of possible phenotypic states of a reconstructed biological network in a given situation. Furthermore, constraint-based reconstruction and analysis tools theoretically study the distribution of metabolic flux in genome-scale network-based models.

The omics (transcriptomics, proteomics, and metabolomics) and bioinformatics based on GSMs are integral parts of systems biology along with bioinformatics with GSMs that help to understand a complex metabolic network by allowing a holistic picture of the dynamic system having different levels of biological organization interacting with the external environment for phenotypic expression [21]. In addition, multi-omics datasets can further help to identify candidate genes, and specific promoters involved in the biochemistry of NUE and stress adaptation in C4 millets [5, 21]. The obvious choice for an omics experiment is to supplement millet with varying concentrations of N. Such an analysis would identify which genes are differentially regulated together with the enzymes of N metabolism when there is too much or too little of N. These genes usually provide important clues and reveal both negative and positive regulators of a pathway.

Since C/N metabolism intersects closely in photoautotrophic organisms, these experiments can also be performed under varying concentrations of C. Since the carboxylation enzymes of C4 plants like millets store less N and are likely less affected by N limitation, their C/N metabolism is expected to be distinct from that of C3 plants. It may be essential to incorporate mutants (knockout and overexpression) of regulatory genes involved in the balance of C/N metabolism in systems biology investigations on millets (Figure 2) [10]. This will provide additional information on the regulatory mechanism that can be used in further manipulation of engineered major crops to optimize the target pathways and maximize their yield. For example, including the mutant of AtGLR1.1 (a putative glutamate receptor) gene/orthologues

in systems biology of millets could be useful in terms of exploring C/N metabolism as its function as a regulator of C/N metabolism in *Arabidopsis thaliana* has been proposed earlier [22]. In contrast to plant C/N metabolism, there is a substantial amount of knowledge available on cyanobacterial C/N metabolism that can be used to inspire engineering ideas (Box 1).

The systems biology of millets should also aim to study the metabolic performance of crops grown under different N forms such as nitrate and the N metabolites derived from nitrate reduction and assimilation pathways (ammonia, glutamate, etc.). These investigations may aid in differentiating the overall impact of nitrate from the signaling impact of metabolites produced after nitrate reduction and assimilation. Further, including enzyme inhibitors such as MSX (L-methionine-sulfoximine (MSX), an inhibitor of glutamine synthetase that mimics the metabolic condition of N limitation, could further distinguish the pathways responsive only to N starvation from the other responses (e.g., redox imbalance) that are triggered indirectly by lack of N in cells [10]. Finally, a systems and omics level comparison of C4 millets with their closest C3 crops could be one promising way to better understand C4 metabolism in general together with the N metabolism of millets [23,24].

Systems metabolic engineering of nitrogen metabolism in millets

The effective metabolic engineering of NUE in major crops will require synthetic biology methods to deliver key genes (or genes for rate-limiting enzymes) or the entire pathway from millets to major crops (Figure 1). Then, using these tools, the transferred pathways can be further optimized in the engineered major crop to attain the maximum productivity under N-limited conditions. The three main components of metabolic engineering strategies are the introduction of biosynthetic genes, fine-tuning of target gene expression, and improvement or rerouting of intracellular metabolic flux (rate of turnover of metabolites through a metabolic pathway). The increase in the expression of the genes that encode precursors and rate-limiting enzymes can ensure the abundance of precursor molecules that could increase metabolic flux towards the production of the target metabolite (that confers tolerance to N or C limitation). Further, the deletion of competitive sub-branches in the target pathway should be done to avoid the cellular resources and intermediates mobilizing towards the other less important metabolites. In addition, the overexpression of transcription factors to simultaneously activate the multiple pathways towards the enhanced synthesis of metabolites could be considered. Furthermore, these approaches, when used in combination, could maximize the crop productivity. With the help of **CRISPR/Cas** technologies, such time-consuming multiplexed gene manipulation is now possible in both labor-and time-effective manners [24]. When overexpressed in rice and *Arabidopsis*, a plant-specific Dof1 (DNA binding with

one finger) transcription factor activates the C4-related PEPc (phosphoenolpyruvate carboxylase) enzyme with increased amino acid content, increased C skeleton, and a decrease in glucose levels [25]. A considerable variation in amino acid was seen in *Arabidopsis* plants that overexpressed Dof1 in N-limited conditions, suggesting that Dof1 may play a key role in plant NUE [26]. According to another study, Dof1 appears to enhance NUE uptake and assimilation in low N environments [27]. Besides Dof1, the PII enzyme, a protein that senses and regulates N, is yet another possible target for engineering NUE in crops (Box 1). The constitutive overexpression of PII-like protein/homologue GLB1 reduces the capacity to recognize and mobilize glutamine [28]. Additional research has indicated that nitrite uptake into plant chloroplasts increased in GLB-PII knockout mutants, indicating that PII may be a limiting element in N uptake and assimilation [29]. The systems biology of millets may provide insight into the global profiles of genes and metabolites activated in them, including the status of PII or PII-like proteins and Dof1 proteins in cells, and forecast a relationship with other enzymes that upregulate or downregulate with these proteins simultaneously. Further research, including systems analysis of mutants of these genes under both N-repleted and N-depleted conditions, could shed light on their regulators and interacting pathways. Table 2 provides a list of additional genes that could be suitable for understanding and improving NUE in transgenic crops.

An emerging paradigm in metabolic engineering is the Design-Build-Test-Learn (DBTL) cycle, presenting another opportunity for successful implementation of systems metabolic engineering strategy to create major crops with higher NUE [30]. This strategy represents a methodical and effective strategy for strain development efforts in microbiology but is underused in plant research [30, 31]. Growing interest in the DBTL cycle for metabolic engineering is largely a result of advancements made in synthetic biology (such as DNA synthesis, genome editing, and synthetic biology tools), omics technology, and machine learning methods [32, 33]. In the DBTL cycle, genetic constructs are designed and constructed in microbial hosts using synthetic biology, and the knowledge obtained from omics technologies during the cycle's test phase is then transferred to learning processes. In order to promote the engineering biology, aim for additional strain creation and optimization, what is learned (such as constraint-based FBA models) is then fed back to new cycles of design. This makes it easier to quickly optimize microbial strains for the production of any desired chemical compound [33]. Since mathematical models (of the engineered bioproduct, route, biological system, or biome) are only as good as their assumptions, the learning process is arguably the weakest link in the DBTL cycle workflow [30]. As a result, in order to enhance training models and ensure greater accuracy and robustness of the learning process, big and high-quality omics data sets are required. The systems analysis can particularly

improve the design phase of the DBTL cycle in the implementation of the NUE mechanism of millets in major crops, as it not only involves computer-based models but also practical data on metabolic flux.

Challenges

N metabolism and NUE are complex multigenic traits, and therefore their improvement becomes difficult, particularly in C4 crops. Furthermore, under natural field conditions, the interactions between N, water, and other nutrients complicate the assessment of NUE, so it is often dissected into subcomponent traits. Identification and functional characterization of candidate NUE genes existing in pathways relating to N metabolism, NUE and C/N signaling and regulation in these neglected crops is somewhat difficult due to a scarcity of forward and reverse genetics tools. Therefore, a systems biology approach is essential to depict the regulation of N uptake, N assimilation, and N recycling in a dynamic and integrated manner. To effectively implement the systems biology approaches in millets, a variety of mutants would be required as a control and to verify novel phenotypes. This requires the establishment of genetic engineering methodologies and synthetic biology tools for millets. However, with a limited understanding of the biology and metabolism of millets, it remains challenging to design and build genetic manipulation strategies for their effective engineering. Since, metabolic pathways in C4 crops span different organelles and compartments with a series of organellar transporters, another key challenge includes the characterization of intracellular transporters and their appropriate inclusion in context-specific plant metabolic models [15, 34, 35]. Metabolic flux profiling can also be influenced by a high level of compartmentalization, which implies that a particular label and associated metabolite pool can originate from various, spatially separated organelles/compartments. Currently, the systems biology of millets is lagging behind also due to the lack of their reference genomes, as they have remained overlooked in plant research. Although the draft/reference genome sequence of a number of millet species has been released, as of now the fully annotated genome sequence is available only for foxtail millet [15]. Furthermore, even though large-scale metabolic models for several plants have been recreated, no N metabolism specific GSM of millet has been presented so far. GSM is essential for determining whether specific genetic manipulation aimed at shifting metabolic flux to desired end-products is effective. The rectification of above-mentioned issues might be fundamental in millet-inspired systems metabolic engineering of NUE in major crops.

Concluding remarks

Millets are usually more adapted to low-input farming, and as many of them such as millets have a C4 metabolism, they are more productive than C3 plants in terms of biomass production and photosynthetic rate per unit of N. This efficient metabolism of N uptake and utilization has remained obscured due to a lack of research focus and breeding programs in millets. A system metabolic engineering approach is therefore required to unravel the key mechanism behind their high NUE (Figure 1). Moreover, analysis of NUE phenotypes and their progression during growth and development under varying N fertilization in the natural field is urgently required for an accurate assessment of NUE. To do so, the systems biology approach (metabolic flux modelling and omics) can be highly useful to delineate the metabolic interactions of N utilization pathways at the global scale and in the identification of genes for key rate-limiting enzymes, regulators, and pathways. These mechanisms, once identified, could be used in the metabolic engineering of major crops that are a more popular and globally consumed source of food. These ideas may face challenges due to the lack of fully annotated genomes of millets, lack of defined genetic engineering tools specific to millets, and the presence of different compartments and organelles being a C4 plant. Furthermore, large data sets are needed, and several pertinent questions (see Outstanding questions) need to be addressed to create holistic interaction maps, linking genes to each other and to their biological functions. These issues could be targeted in the future with the help of standardization and optimization of metabolic flux models and establishing genetic engineering protocols for millets. A new era of systems metabolic engineering of major staple crops using genes from millets may result in a more rational management of their key genetic resources, ensure the best adapted genetic diversity, and leverage this crop diversity into breeding programs, paving the way towards more resilient, resource-efficient, and diverse agricultural ecosystems.

Acknowledgements

This work is supported by a Ramanujan Fellowship (file no. RJF/2020/000043) awarded to P.K.B. by Science and Engineering Research Board (SERB), Department of Science and Technology (DST), New Delhi, India.

Declaration of interests

None are declared.

References

1 Erisman, J.W. et al. (2008) How a century of ammonia synthesis changed the world. Nat. Geosci. 1, 636-639.

2 Erisman, J.W. et al. (2011) Reactive nitrogen in the environment and its effect on climate change. Curr. Opin. Environ. Sustain. 3, 281–290

3 Choi, K.R. et al. Systems metabolic engineering strategies: Integrating systems and synthetic biology with metabolic engineering. Trends Biotechnol. 37, 817–837

4 Ye, C.-Y. and Fan, L. (2021) Orphan crops and their wild relatives in the genomic era. Mol. Plant 14, 27–39

5 Babele, P.K. et al. (2022) Mainstreaming orphan millets for advancing climate smart agriculture to secure nutrition and health. Front. Plant Sci. 2806

6 Brown, R.H. (1978) A difference in N use efficiency in C3 and C4 plants and its implications in adaptation and evolution. Crop Sci. 18, 93–98

7 Makino, A. et al. (2003) Differences between Maize and Rice in N-use efficiency for photosynthesis and protein allocation. Plant Cell Physiol. 44, 952–956

8 Evans, J.R. and Von Caemmerer, S. (1996) Carbon dioxide diffusion inside leaves. Plant Physiol. 110, 339–346

9 Ghannoum, O. (2009) C4 photosynthesis and water stress. Ann. Bot. 103, 635–644

10 Kumar, A., Gupta, N., Gupta, A. K., & Gaur, V. S. (2009). Identification of biomarker for determining genotypic potential of nitrogen-use-efficiency and optimization of the nitrogen inputs in crop plants. J. Crop Sci. Botechnol. 12(4), 183-194.

11 Kumar, R., Taware, R., Gaur, V. S., Guru, S. K., & Kumar, A. (2009). Influence of nitrogen on the expression of TaDof1 transcription factor in wheat and its relationship with photo synthetic and ammonium assimilating efficiency. Mol. Biol. Rep. 36(8), 2209-2220.

12 Shih, M.L. and Morgan, J.A. Metabolic flux analysis of secondary metabolism in plants. Metab. Eng. Commun. 10, e00123

13 Xu, Chengjie, Mingzhao Luo, Xianjun Sun, Jiji Yan, Huawei Shi, Huishu Yan, Rongyue Yan et al. (2022) SiMYB19 from foxtail millet (*Setaria italica*) confers transgenic rice tolerance to high salt stress in the field." Int. J. Mol. Sci. 23(2), 756.

14 Gu, C. et al. (2019) Current status and applications of genome-scale metabolic models. Genome biol. 20, 1-18.

15 Allen, D.K. et al. (2009) Metabolic flux analysis in plants: Coping with complexity. Plant Cell Environ. 32, 1241–1257

16 Young, J.D. et al. (2011) Mapping photoautotrophic metabolism with isotopically nonstationary 13C flux analysis. Metab. Eng. 13, 656–665
17 Babele, P.K. and Young, J.D. (2020) Applications of stable isotope-based metabolomics and fluxomics toward synthetic biology of cyanobacteria. Wiley Interdiscip. Rev. Syst. Biol. Med. 12, e1472.
18 Ma, F. et al. (2014) Isotopically nonstationary 13C flux analysis of changes in Arabidopsis thaliana leaf metabolism due to high light acclimation. Proc. Natl. Acad. Sci. U. S. A. 111, 16967–16972
19 Young, J.D. (2014) INCA: a computational platform for isotopically non-stationary metabolic flux analysis. Bioinformatics 30, 1333–1335
20 Shaw, R. and Maurice Cheung, C.Y. (2019) A mass and charge balanced metabolic model of Setaria viridis revealed mechanisms of proton balancing in C4 plants. BMC Bioinformatics 20, 1–11
21 Pazhamala, L.T. et al. (2021) Systems biology for crop improvement. Plant Genome 14, e20098
22 Kang, J. and Turano, F.J. (2003) The putative glutamate receptor 1.1 (AtGLR1.1) functions as a regulator of carbon and nitrogen metabolism in Arabidopsis thaliana. Proc. Natl. Acad. Sci. U. S. A. 100, 6872
23 Wang, C. et al. (2012) Systematic comparison of C3 and C4 plants based on metabolic network analysis. BMC Syst. Biol. 2012 62 6, 1–14
24 Blätke, M.A. and Bräutigam, A. (2019) Evolution of C4 photosynthesis predicted by constraint-based modelling. Elife, 8,
25 Chen, K. et al. (2019) CRISPR/Cas genome editing and precision plant breeding in agriculture. Annu. Rev. Plant Biol. 70, 667–697
26 Yanagisawa, S. et al. (2004) Metabolic engineering with Dof1 transcription factor in plants: Improved nitrogen assimilation and growth under low-nitrogen conditions. Proc. Natl. Acad. Sci. U. S. A. 101, 7833–7838
27 Gupta, S. et al. (2014) Fluctuation of Dof1/Dof2 expression ratio under the influence of varying nitrogen and light conditions: involvement in differential regulation of nitrogen metabolism in two genotypes of finger millet (Eleusine coracana L.). Gene 546, 327–335
28 Kurai, T. et al. (2011) Introduction of the ZmDof1 gene into rice enhances carbon and nitrogen assimilation under low-nitrogen conditions. Plant Biotechnol. J. 9, 826–837
29 Hsieh, M.H. et al. (1998) A PII-like protein in Arabidopsis: Putative role in nitrogen sensing. Proc. Natl. Acad. Sci. U. S. A. 95, 13965–13970

30 Ferrario-Méry, S. et al. (2008) Chloroplast nitrite uptake is enhanced in Arabidopsis PII mutants. FEBS Lett. 582, 1061–1066
31 Pouvreau, B. et al. From plant metabolic engineering to plant synthetic biology: The evolution of the design/build/test/learn cycle. Plant Sci. 273, 3–12
32 Lawson, C.E. et al. (2019) Common principles and best practices for engineering microbiomes. Nat. Rev. Microbiol. 17, 725–741
33 Zhu, Q. et al. (2020) Plant synthetic metabolic engineering for enhancing crop nutritional quality. Plant Commun. 1, 100017
34 Carbonell, P. et al. (2018) An automated design-build-test-learn pipeline for enhanced microbial production of fine chemicals. Commun. Biol. 11, 1–10
35 Cui, H. (2021) Challenges and approaches to crop improvement through C3-to-C4 engineering. Front. Plant Sci. 12, 1851
36 Baslam, M. et al. (2021) Recent advances in carbon and nitrogen metabolism in C3 plants. Int. J. Mol. Sci. 22, 1–39
37 Araus, V. et al. (2020) A balancing act: how plants integrate nitrogen and water signals. J. Exp. Bot. 71, 4442–4451
38 Forchhammer, K. and Selim, K.A. (2020) Carbon/nitrogen homeostasis control in cyanobacteria. FEMS Microbiol. Rev. 44, 33–53
39 Forchhammer, K. et al. (2022) New views on PII signaling: from nitrogen sensing to global metabolic control. Trends Microbiol. 30, 722-735
40 Selim, K.A. et al. (2019) Tuning the in vitro sensing and signaling properties of cyanobacterial PII protein by mutation of key residues. Sci. Rep, 91, 1–11
41 Selim, K.A. et al. (2020) From cyanobacteria to archaeplastida: |New evolutionary insights into PII signalling in the plant kingdom. New Phytol. 227, 722–731
42 Selim, K.A. et al. (2018) PII-like signaling protein SbtB links cAMP sensing with cyanobacterial inorganic carbon response. Proc. Natl. Acad. Sci. U. S. A. 115, e4861–e4869
43 Mantovani, O. et al. (2022) The impact of the cyanobacterial carbon-regulator protein SbtB and of the second messengers cAMP and c-di-AMP on CO2-dependent gene expression. New Phytol. 234, 1801-1816.
44 Orthwein, T. et al. (2021) The novel PII-interactor PirC identifies phosphoglycerate mutase as key control point of carbon storage metabolism in cyanobacteria. Proc. Natl. Acad. Sci. U. S. A. 118, e2019988118

45 Selim, K.A. et al. (2021) Diurnal metabolic control in cyanobacteria requires perception of second messenger signaling molecule c-di-AMP by the carbon control protein SbtB. Sci. Adv. 7,
46 Scholl, J. et al. (2020) Phosphoenolpyruvate carboxylase from the cyanobacterium Synechocystis sp. PCC 6803 is under global metabolic control by PII signaling. Mol. Microbiol. 114, 292–307
47 Hauf, W. et al. (2016) Interaction of the nitrogen regulatory protein GlnB (PII) with biotin carboxyl carrier protein (BCCP) controls acetyl-CoA levels in the cyanobacterium Synechocystis sp. PCC 6803. Front. Microbiol. 7, 1700
48 Bourrellier, A.B.F. et al. (2010) Chloroplast acetyl-CoA carboxylase activity is 2-oxoglutarate-regulated by interaction of PII with the biotin carboxyl carrier subunit. Proc. Natl. Acad. Sci. U. S. A. 107, 502–507
49 Bolay, P. et al. (2021) The novel PII-interacting protein PirA controls flux into the cyanobacterial ornithine-ammonia cycle. MBio 12, e00229-21
50 Rajendran, C. et al. (2011) Crystal structure of the GlnZ-DraG complex reveals a different form of PII-target interaction. Proc. Natl. Acad. Sci. U. S. A. 108, 18972–18976
51 Nordlund, S. (2015) Metabolic Regulation of Nitrogenase Activity in Rhodospirillum rubrum: The Role of PII Proteins and Membrane Sequestration. Biol. Nitrogen Fixat. 1–2, 131–138
52 Dal'Molin, C.G. de O. et al. (2010) C4GEM, a genome-scale metabolic model to study C4 plant metabolism. Plant Physiol. 154, 1871–1885
53 Saha, R. et al. (2011) Zea mays iRS1563: A comprehensive genome-scale metabolic reconstruction of maize metabolism. PLoS One 6, e21784
54 Simons, M. et al. (2014) Assessing the metabolic impact of nitrogen availability using a compartmentalized maize leaf genome-scale model. Plant Physiol. 166, 1659–1674
55 diCenzo, G.C. et al. (2020) Genome-scale metabolic reconstruction of the symbiosis between a leguminous plant and a nitrogen-fixing bacterium. Nat. Commun. 11, 1–11
56 Moreira, T.B. et al. A genome-scale metabolic model of soybean (glycine max) highlights metabolic fluxes in seedlings. Plant physiol., 180, 1912-1929.
57 Grafahrend-Belau, E. et al. (2009) Flux balance analysis of barley seeds: A computational approach to study systemic properties of central metabolism. Plant Physiol. 149, 585–598
58 Schulte, C.C.M. et al. (2021) Metabolic control of nitrogen fixation in rhizobium-legume symbioses. Sci. Adv. 7, eabh2433
59 Zhang, J. et al. (2021) Increasing yield potential through manipulating of an ARE1 ortholog related to nitrogen use efficiency in wheat by CRISPR/Cas9. J. Integr. Plant Biol. 63, 1649–1663

60 Chen, K.-E. et al. (2020) Improving nitrogen use efficiency by manipulating nitrate remobilization in plants. Nat. Plants, 69, 1126–1135
61 Chen, J. et al. (2017) pOsNAR2.1: OsNAR2.1 expression enhances nitrogen uptake efficiency and grain yield in transgenic rice plants. Plant Biotechnol. J. 15, 1273
62 Yuan, L. et al. (2007) Nitrogen-dependent posttranscriptional regulation of the ammonium transporter AtAMT1;1. Plant Physiol. 143, 732–744
63 Hirner, A. et al. (2006) Arabidopsis LHT1 is a high-affinity transporter for cellular amino acid uptake in both root epidermis and leaf mesophyll. Plant Cell 18, 1931–1946
64 Schofield, R.A. et al. (2009) Over-expression of STP13, a hexose transporter, improves plant growth and nitrogen use in Arabidopsis thaliana seedlings. Plant. Cell Environ. 32, 271–285
65 Perchlik, M. and Tegeder, M. (2017) Improving plant nitrogen use efficiency through alteration of amino acid transport processes. Plant Physiol. 175, 235–247
66 AK, S. et al. (2008) Genetic engineering of improved nitrogen use efficiency in rice by the tissue-specific expression of alanine aminotransferase. Plant Biotechnol. J. 6, 722–732
67 Tiong, J. et al. (2021) Improving nitrogen use efficiency through overexpression of alanine aminotransferase in rice, wheat, and barley. Front. Plant Sci. 29
68 Lee, S. et al. (2020) OsASN1 Overexpression in rice increases grain protein content and yield under nitrogen-limiting conditions. Plant Cell Physiol. 61, 1309–1320
69 Wu, D. et al. (2021) Increased glutamine synthetase by overexpression of TaGS1 improves grain yield and nitrogen use efficiency in rice. Plant Physiol. Biochem. PPB 169, 259–268
70 Zeng, D.D. et al. (2017) The ferredoxin-dependent glutamate synthase (OsFd-GOGAT) participates in leaf senescence and the nitrogen remobilization in rice. Mol. Genet. Genomics 292, 385–395
71 Sun, H. et al. (2014) Heterotrimeric G proteins regulate nitrogen-use efficiency in rice. Nat. Genet. 46, 652–656
72 Han, M. et al. (2016) Identification of nitrogen use efficiency genes in barley: searching for qtls controlling complex physiological traits. Front. Plant Sci. 0, 1587
73 Bi, Y.M. et al. (2009) Increased nitrogen-use efficiency in transgenic rice plants over-expressing a nitrogen-responsive early nodulin gene identified from rice expression profiling. Plant. Cell Environ. 32, 1749–1760
74 Rubio-Wilhelmi, M.D.M. et al. (2011) Cytokinin-dependent improvement in transgenic PSARK::IPT tobacco under nitrogen deficiency. J. Agric. Food Chem. 59, 10491–10495

75 Chiasson, D.M. et al. (2014) Soybean SAT1 (Symbiotic Ammonium Transporter 1) encodes a bHLH transcription factor involved in nodule growth and NH4+ transport. Proc. Natl. Acad. Sci. U. S. A. 111, 4814–4819

76 He, X. et al. (2015) The nitrate-inducible NAC transcription factor tanac2-5a controls nitrate response and increases wheat yield. Plant Physiol. 169, 1991

77 Li, W. et al. (2020) A wheat transcription factor positively sets seed vigour by regulating the grain nitrate signal. New Phytol. 225, 1667–1680

Glossary

C3 plants: These plants use the enzyme RuBisCO in relatively inefficient ways, to fix CO_2 and obtain the 3-carbon organic compound, 3-phosphoglycerate (3PGA).

C4 plants: In the C4 plants, PEPc links CO_2 to a C3 acid, phosphoenolpyruvate (PEP), forming a C4 acid, oxaloacetate (OAA).

CO_2 concentration mechanism: In order to decrease Rubisco's oxygenase activity and, consequently, the rate of photorespiration, C4 photosynthesis concentration CO_2 near Rubisco

Constraints based reconstruction and analysis: It provides a molecular mechanistic framework for integrative analysis of experimental molecular systems biology data and quantitative prediction of physiochemically and biochemically feasible phenotypic states.

CRISPR/Cas: It refers to Clustered Regularly Interspaced Short Palindromic Repeats, which are the hallmark of a bacterial defense system that forms the basis for a robust genome editing technology.

Flux balance analysis: It is a mathematical approach for analyzing the flow of metabolites through a metabolic network.

Fluxomics: It is used to quantify and assess the rates of reactions (fluxes) for a network of metabolic reactions in an organism.

Genome-scale metabolic model: It is an *in silico* metabolic model comprising hundreds or thousands of chemical reactions that constitute the metabolic inventory of a cell, tissue, or organism.

Major crops: Unlike orphan/minor crops, major crops (e.g., wheat, rice etc.) are widely consumed sources of food and have been the focus of plant breeding and biotechnology research globally.

Metabolic flux analysis: It provides a quantitative description of the flow of metabolites within a biological network and relies on ^{13}C or other isotopes to track, or enrich, metabolites according to biochemical fluxes and pathways.

Metabolic flux: The rate at which molecules move through a metabolic pathway is known as the metabolic flux.

Metabolically engineered cultivars: The genetically modified plants with rationally targeted and altered metabolic pathways. These manipulations usually aid in the better utilization of cellular resources for the production of target metabolites or desired phenotypes.

Metabolomics: The broad study of metabolites found in cells or organisms is known as metabolomics. The metabolome refers to all of these small molecules as well as how they interact with one another inside a biological system.

NUE: It is a parameter to describe the ability of a crop to uptake, assimilate, and utilize nitrogen.

Orphan crops: Crops that are often not traded internationally and farmed only locally.

Photorespiration: In photorespiration, the enzyme RuBisCO oxygenates RuBP (Ribulose 1,5-bisphosphate, a CO_2 acceptor), in which some of the energy from photosynthesis is lost.

Systems biology: It is an integrative and holistic approach including computation, mathematical analysis, and modeling of complex biological systems with a focus on the complex relationships between biological entities.

Systems metabolic engineering: A multidisciplinary approach that combines traditional metabolic engineering with systems biology, synthetic biology, and evolutionary engineering.

Box 1. Regulation of C/N metabolism by cyanobacterial PII protein(s): clues from prokaryotic ancestors for engineering eukaryotic successors

In photoautotrophs, the C/N assimilation reactions *via* the Calvin-Benson and Glutamine-synthetase-glutamate synthetase (GS-GOGAT) cycles require tight regulation and constant sensing of the quantity and quality of carbon and nitrogen availability [36]. Generally, carbon and nitrogen metabolisms are coordinated through a complex crosstalk among different input signals [37]. Therefore, cyanobacteria and higher plants have evolved a sophisticated signal transduction network, which rely on the signal-transduction proteins of the PII superfamily for controlling C/N homeostasis [37, 38]. Therefore, in order

to understand the systems biology of orphan and staple crops and unravel the targets for engineering NUE, PII proteins and their interacting network represent an important candidate. PII proteins sense the energy, carbon, and nitrogen status of the cell through the binding of small effector molecules via the binding of ATP or ADP competitively to the same binding site and the TCA cycle intermediate 2-oxoglutrate (2-OG) [39, 40]. Moreover, a new class of PII-like proteins has been discovered and especially the PII-like protein, SbtB, was shown to sense second messengers cAMP and c-di-AMP, which reflect the carbon status and the diurnal status of the cell, respectively [41-43].

Cyanobacteria harbor three bicarbonate uptake systems (SbtA, BicA, and BCT1-complex) and three nitrogen uptake systems (nitrate Nrt-complex, urea Urt-complex, and ammonium transporter AmtB), which are regulated either by PII or SbtB [41-43]. Both PII and SbtB complex with AmtB and SbtA, respectively, to regulate their transport activity [38, 41]. The accumulated bicarbonate is dehydrated to CO_2 by carbonic anhydrase and then fixed by the RuBisCO enzyme to 3-phosphoglycerate (3-PGA) initiating the Calvin-Bensson-Bassham cycle [37, 42]. 3-PGA, the first stable carbon-fixation product, is now directed into different routes to fill glycogen anabolism and TCA cycle via PGAM (phosphoglycerate mutase) activity. PGAM is controlled positively by PII *via* sequestering the PGAM inhibitor PirC [44], while glycogen anabolism is regulated by SbtB *via* controlling the glycogen-branching enzyme GlgB [45]. PII senses the changes in the 2-OG levels, which is of particular interest as 2-OG reports the changes in the C/N balance of CO_2 fixation and ammonium assimilation: the 2-OG levels are depleted by ammonium assimilation reactions (GS-GOGAT cycle and glutamate dehydrogenase) and refilled by carbon flux originating from carbon-fixation reactions through 3-PGA [37, 38]. The phosphoenol pyruvate carboxylase (PEPC) reaction plays a major role for the carbon flux into the oxidative branch of the TCA cycle and 2-OG synthesis [37, 38]. The PEPC reaction is also controlled positively by PII [46]. Through binding to the biotin carboxylase carrier protein subunit of acetyl-CoA carboxylase (ACCase), PII controls as well the fatty acid biosynthesis [37, 38], which seems a highly conserved target for PII between cyanobacteria [47] and higher plants (as shown for *Arabidopsis thaliana*) [48]. Therefore, it's conceivable to assume that PII superfamily controls the carbon flow in cell via regulating SbtA, PEPC, PGAM and ACCase. Glutamate is the primary nitrogen donor for anabolic reactions, with arginine biosynthesis being of particular importance, whereas PII regulates arginine biosynthesis positively via direct interaction with N-acetyl glutamate kinase and negatively via PirA [37, 38, 49]. Notably, the NAGK is the most highly conserved target for PII between Archaeplastida and cyanobacteria [48]. Finally, in cyanobacteria, PII controls the transcription of nitrogen related genes via sequestering the coactivator (PipX) of the master nitrogen transcription factor

NtcA, keeping it in an inactive state [37, 39]. It is noteworthy to mention that in the diazotrophs, PII plays a key role in the regulation of nitrogenase activity via controlling the nitrogenase regulatory proteins (e.g., NifA, DraG and DraT) [50, 51]. Thus, it is apparent that the signaling proteins of PII superfamily play fundamental roles in controlling the metabolic flow in the cell *via* controlling different targets through sensing the energy, carbon, and nitrogen state, which can be useful in designing engineering strategies for transgenic plants with higher NUE.

Figure 1. Key figure. **Systems metabolic engineering strategies and procedures for improving NUE in crops.** (**A**) The process of GSM reconstruction and validation involves creating an initial model, curating the model, and then rendering the species and context specific model ready for FBA. The first step in curating a plant GSM is to identify all the genes present in the target crop species. There are a number of automated software for genome annotation available for this process. The draft reconstruction is then refined by the user in an iterative manner through an exhaustive review of each reaction, metabolite, and gene in the network. (**B**) Comparison of the genetic and metabolic architecture of C3 (major crops) and C4 crops (millets), and integration of multi-omics (transcriptome, proteome, and metabolome) data and bioinformatics with systems biology tools is essential to understand the mechanism of superior traits in C4 crops. These techniques together present a holistic picture of the dynamic system with the different levels of biological organization interacting with the external environment for phenotypic expression. (**C**) Reconstructed GSM is then transformed into a mathematical structure (stoichiometric matrix), thereafter an objective function is defined, constraints are set and gap filling (addition of reactions to GSM to permit those models to run correctly) was performed to account for specific growth conditions (normal *vs* stress). (**D**) MFA requires estimation of metabolites which is done by nuclear magnetic resonance (NMR) and mass spectroscopies (MS), Modern MFA relies on two possible approaches: steady-state ^{13}C-MFA and INST-MFA where metabolites are labelled with stable isotopes (^{13}C, ^{15}N). (**E**) The outcome of GSM and FBA analysis is used to estimate the fluxes through large metabolic pathways and then generate a simplified model for NUE. MFA quantifies the flow of metabolites through a core metabolic pathway, yielding flux maps that can aid engineering efforts by explaining phenotypes in detail. **Abbreviations**: NUpE- N uptake efficiency, NUpT- total N uptake, KEGG- Kyoto Encyclopedia of Genes and Genomes, RNASeq- ribonucleic acid sequencing, V- reaction rate, SSR- sum of squared residual

Figure 2. Schematic model of nitrogen transport, sensing and metabolic basis of C/N balance in plants. The figure illustrates the difference between C3 and C4 metabolism and the cellular compartments they operate in. The C3 pathway takes place in mesophyll cells and C4 in both mesophyll and bundle sheath cells. The CO_2 acceptor molecule in the C3 pathway is RuBP, whereas that in the C4 pathway is PEP. The first stable products in the C3 and C4 pathways are a three-carbon compound called 3PGA and a four-carbon compound called OAA, respectively. In C3 plants, the photorespiration rate is high and leads to the loss of fixed carbon dioxide, but in C4 plants the photorespiration rate is negligible, which increases the CO_2 fixation rate in C4 plants. C3 plants exhibit a high photorespiratory pathway, which is initiated by the oxygenase activity of RuBisCO to form 2PG. 2PG is then converted to glycolate and then to glycine in both peroxisomes and mitochondria. The carbon metabolic pathway generates energy (ATP) and reduces potential NAD(P)H for nitrogen assimilation. The carbon skeleton part in amino acids also comes from 2-OG of the TCA (Tricarboxylic acid cycle) cycle. Nitrate (NO_3^-) is reduced by nitrate reductase to nitrite (NO_2^-) and further by nitrite reductase to ammonium (NH_4^+). 2OG serves as a C skeleton for the synthesis of Glu by incorporating photorespiratory NH_4^+. NH_4^+ from the primary N assimilation is then incorporated into Glu, resulting in the production of Gln. Glu and Gln donate NH_4^+ used for the synthesis of all other amino acids, including Asp or Asn, which serve as either an active NH_4^+ donor (Asp) or a transport/storage compound (Asn). In response to the heterogeneity in inorganic nitrogen concentration in the soil, plants have evolved mechanisms to regulate its influx. Plants are able to sense $NO3^-$ and $NH4^+$ in their environment through transceptors (transporter/receptor) that activate several downstream signaling cascades such as mitogen-activated protein kinases (MAPK, MAPKK), and transcription factors (TFs). Inorganic nitrogen-dependent activation of cell signaling cascades modulates expression of nitrogen metabolism (NM) genes in the nucleus. These molecular events are also regulated by epigenetic markers such as DNA methylation, histone modifications, and expression of noncoding RNAs. **Abbreviations**: NRT- nitrate transporter; AMT- ammonia transporter; NR- nitrate reductase; NiR- nitrite reductase; 2OG- 2-oxoglutarate; 3PGA- 3-phosphoglyceric acid; 2PG- 2-phosphoglycerate; (NiR), G3P- glyceraldehyde-3-phosphate; Asn-asparagine; Cit- citrate; GS- glutamine synthetase; Gly- glycine; GOGAT- glutamate synthase; Glu- glutamate; Gln- Glutamine; Hpyr- hydroxypyruvate; Mal- malate; OAA-oxaloacetic acid; PEP- phosphoenol pyruvate; Pyr- pyruvate; PPP- pentose phosphate pathway; RuBP- ribulose 1,5-bisphosphate; Ser- serine; Suc- succinate; TF- transcription factor.

581 **Table 1.** Summary of important and most recent GSM models, FBA, MFA and CRISPR/Cas-based
582 studies on model plants and crops for understanding and engineering nitrogen metabolism.
583

Strategies	Model name/ target plant	Approach and applications	Refs
GSM	C4GEM *Arabidopsis*, Sorghum, Maize, Sugarcane	1. First large-scale metabolic model for C4 plants, encapsulate metabolic interactions between two different cell types. 2. Extension of AraGEM with addition of 588 unique reaction, 1755 metabolites, 83 inter-organelle transporters and 29 external transporters.	[52]
	*i*RS1563/ *Zea mays*	1. Contains 1,563 genes and 1,825 metabolites involved in 1,985 primary and secondary maize metabolism reactions. 2. Approximately 42% of the reactions have direct literature evidence for the participation of the reaction in maize. 3. Maize C4GEM (*Zea mays* iRS1563) contains 674 metabolites and 893 reactions that are not accounted for. 4. All reactions are elementally and charged balanced and localized into six different compartments (i.e., cytoplasm, mitochondrion, plastid, peroxisome, vacuole, and extracellular).	[53]
	Maize *(Zea mays)*	1. A second-generation GSM model for the maize leaf was created to capture C4 carbon fixation and investigate nitrogen (N) assimilation by modeling the interactions between the bundle sheath and mesophyll cells. 2. Updated model spans 5,824 genes, 8,525 reactions, and 9,153 metabolites, an increase of approximately 4 times the size of the earlier iRS1563 model. 3. Transcriptomic and proteomic data have also been used to introduce regulatory constraints in the model to simulate an N-limited condition and glutamine synthetase deficient mutants. 4. Model-predicted results achieved 90% accuracy when comparing the wild type grown under an N-complete condition with the wild type grown under an N-deficient condition.	[54]
FBA	*Medicago truncatula*	1. Reconstruction and modelling of a genome-scale metabolic network of *Medicago truncatula* (plant) nodulated by *Sinorhizobium meliloti* (bacterium).	[55]

		2. Reconstructed nodule tissue contains five spatially distinct developmental zones and encompasses the metabolism of both the plant and the bacterium. 3. Revealed that the metabolic costs associated with symbiotic nitrogen fixation are related to nitrogenase activity and increasing N_2-fixation efficiency is associated with diminishing returns in terms of plant growth. 4. Revealed that differentiating bacteroid have access to sugars as major carbon sources, ammonium is the main nitrogen export product of N_2-fixing bacteria, and N_2 fixation depends on proton transfer from the plant cytoplasm to the bacteria through acidification of the peribacteroid space.	
	Soybean (*Glycine max*)	1. A genome-scale stoichiometric model for this important crop plant and then adapting the model to reflect metabolism in the cotyledons and hypocotyl/root axis (HRA) and provided new insight into well-characterized metabolic processes. 2. FBA analysis of seedling growth and alterations in biomass composition revealed marked differences in metabolism between the HRA, together with shifts in primary metabolism occurring during different periods post-germination. 3. Cotyledons were characterized by the oxidation of fatty acids to supply carbon for the tricarboxylic acid cycle as well as production of sucrose and glutamate for export to the HRA, while the HRA was characterized by the use of a range of imported amino acids in protein synthesis and catabolic processes.	[56]
	Barley (*Hordeum vulgare*)	1. The model for primary metabolism (contains 257 biochemical and transport reactions across four different compartments) was subjected to flux balance analysis to study grain yield and metabolic flux distributions in response to oxygen depletion and enzyme deletion. 2. Simulation results were found to be in good agreement with the main biochemical properties of barley seed storage metabolism. 3. The predicted growth rate and the active metabolic pathway patterns under anoxic, hypoxic, and aerobic conditions predicted by the model were in accordance with published experimental results. In addition, the model predictions gave insight into the potential role of inorganic pyrophosphate metabolism in maintaining seed metabolism under oxygen deprivation.	[57]
MFA	*Maize (Zea mays)*		[54]

		1. The updated model spans 5,824 genes, 8,525 reactions, and 9,153 metabolites, an increase of approximately 4 times the size of the earlier iRS1563 model. 2. Transcriptomic and proteomic data have also been used to introduce regulatory constraints in the model to simulate a nitrogen-limited condition and glutamine synthetase deficient mutants. 3. Model-predicted results achieved 90% accuracy when comparing the wild type grown under a nitrogen-complete condition with the wild type grown under a nitrogen-deficient condition.	
	Rhizobium leguminosarum bv. *viciae*	1. Modeling and ^{13}C metabolic flux analysis indicate that oxygen limitation restricts the decarboxylating arm of the tricarboxylic acid cycle, which in turn limits ammonia assimilation into glutamate. 2. By tightly controlling oxygen supply and providing dicarboxylates as the energy and electron source donors for N_2 fixation, legumes promote ammonia secretion by bacteroids. This is a defining feature of *Rhizobium*-legume symbioses.	[58]
CRISPR-Cas	Wheat (*Triticum aestivum* L.)	1. Using CRISPR/Cas9-mediated targeted mutagenesis, a series of transgene free mutant lines with partial or triple null taare1 alleles were generated. 2. Transgene free mutant lines showed enhanced tolerance to nitrogen starvation and showed delayed senescence and increased grain yield in field conditions.	[59]

Table 2. Important studies describing the metabolic components of NUE and their engineering.

Metabolic components	Gene family	Phenotypic description/functions related to NUE	Genetic engineering/outcome	Refs
Transporters				
NRT	Nitrate transporter, nitrate uptake, nitrate transport	Nitrate content and dry weight increased in shoots	In rice, overexpression of *Os*NRT1.1A resulted in a significant increase in Nitrogen utilization, grain output, and a significant decrease in maturity time.	[60]

NAR	Partner protein of NRT2, (NAR2) activator for NRT2, high-affinity nitrate transporter	Nitrate content and dry weight increased in shoots	Rice yield, NUE, and NO_3^- absorption increased when *Os*NAR2.1 was overexpressed using its native promoter.	[61]
AMT	Ammonium transporter, ammonium uptake, ammonium transport	Increased ammonium uptake and reduced dry weight under high ammonia	*At*AMT1.1 transcript accumulation is nitrogen and organ dependent, implying that mRNA turnover is another mechanism for *At*AMT1.1 regulation in response to plant nitrogen status.	[62]
LHT	Lysine histidine Similar to lysine and Transporter histidine specific transporter	Improved plant growth under low N condition	When inorganic nitrogen is sparse, overexpression of LHT1 enhances amino acid intake by ten times, resulting in enhanced nitrogen utilization efficiency.	[63]
STP13	Hexose transporter	Growth, biomass, and NUE increased by application of exogenous sugar	Transgene analysis of STP13 in tobacco BY-2 suspension cells indicated that its gene product is confined to the plasma membrane (PM). The study suggests that increasing carbon availability can improve a plant's nitrogen use.	[64]
AAP	Amino acid permease	Improved both N uptake and utilization efficiency	AAP1-overexpressing plants showed high NUpE in high N soils, (2) NUtE (nitrogen utilization efficiency) in low N environments, and (3) both NUpE (nitrogen uptake efficiency) and NUtE under moderate N supply, thus demonstrate important physiological plasticity through a flexible response to changing N environments	[65]
Amino Acid Biosynthesis				
alaAT	Alanine amino-transferase	Increase both the percentage N and the plant biomass by improving the N uptake efficiency of the plant.	The study presents a detailed analysis of genetic and metabolic responses to AlaAT overexpression, revealing multiple components and pathways that contribute to the nitrogen-use efficiency.	[66] [67]
AS and ASN	Asparagine synthetase	N content and seed yield at high N and low N input	Overexpression of *Os*ASN1 in seedlings leads to greater nitrogen absorption and assimilation, and improved tolerance to nitrogen deficiency.	[68]

GS	Glutamine synthetase	NUE increased under high N condition	By overexpressing *Ta*GS1, transgenic rice plants with better grain yields than wild-type plants have increased GS activity. The transgenic plant improved root nitrogen acquisition and storage during growth phases, as well as nitrogen remobilization to grains.	[69]
GOGAT	Glutamate synthase	Improved grain filling, total nitrogen content, and dry weight	The gogat1 mutant had a lower seed setting rate, grain protein content was considerably higher, total amino acids in the three leaves and the upmost internode grew considerably throughout the grain-filling stage.	[70]
Nitrate Assimilation				
NR	Nitrate reductase	Nitrate content increased in leaves and high NO emission	A hyperactive chimeric nitrate transporter driven by the NRT1.7 promoter was inserted into the nrt1.7 mutant. Transgenic plants absorbed more nitrate and remobilized more nitrogen in sink tissues.	[60]
NiR	Ferredoxin-Nitrite reductase	NO_2^- assimilation increased		[60]
Signaling and Nitrogen Regulation				
DEP1	G-protein subunit	N uptake, assimilation; grain yield increased at moderate levels of nitrogen input		[71]
SnRK/ SAPK	SNF1-related kinase	Higher N uptake efficiency in overexpressing plant		[72]
ENOD	Early nodulin like protein	Increased total amino acids and nitrogen as well as dry biomass and seed yield	The *Os*ENOD93-1 gene was overexpressed in transgenic rice plants, resulting in increased shoot dry biomass and seed yield, increased total amino acid and nitrogen accumulation in the roots, and increased amino acid content in the xylem sap.	[73]
IPT	Isopentenyl transferase	Gene coding the rate-limiting step in cytokinin (CKs) synthesis.	When subjected to nitrogen deficiency, transgenic tobacco overexpressing IPT maintained the leaf/root ratio, demonstrating a greater NUE and better tobacco leaf quality.	[74]

		Delayed senescence when grown under low nitrogen input		
Transcription factors				
Dof1	DNA-binding one zinc finger	Increased growth, nitrogen assimilation, and enhanced grain production	Dof1 expression resulted in upregulation of genes encoding carbon skeleton-producing enzymes, amino acid content, and a decrease in glucose levels in transgenic *Arabidopsis*.	[25]
SAT1	bHLH transcription factor	Nodulation to improve N fixation and NH_4^+ transport	Since decreased GmbHLHm1 activity limits nodule fitness and development, it is critical for the soybean rhizobium symbiosis.	[75]
NFY	Nuclear factor Y	Increased drought and salinity tolerance and grain yield		
NAC	NAM, ATAF1,2, and CUC2	Enhanced drought resistance; senescence, nutrient remobilization, and grain protein content	In wheat, overexpression of *Ta*NAC2-5A promotes root development and nitrate inflow, increasing the root's ability to acquire nitrogen. Overexpression of *Ta*NAC2-5A increases grain nitrate concentration and seed vigor by binding directly to the promoter of TaNRT2.5-3B.	[76] [77]

587

Outstanding questions

1. To what extent are C and N metabolism in orphan crops interdependent?
2. What are the key spots (genes, enzymes, metabolites) of feedback regulation and crosstalk in nitrogen uptake and utilization of orphan crops?
3. Are NUE-related traits lost in major crops during evolution? Can such traits be restored by reintroducing orphan crop genes into major crops?
4. What is the PII-sensing mechanism in orphan millets and to what extent is it affected by an imbalance in carbon metabolism?

Figure1

Click here to access/download;Key Figure;Figure1.jpg

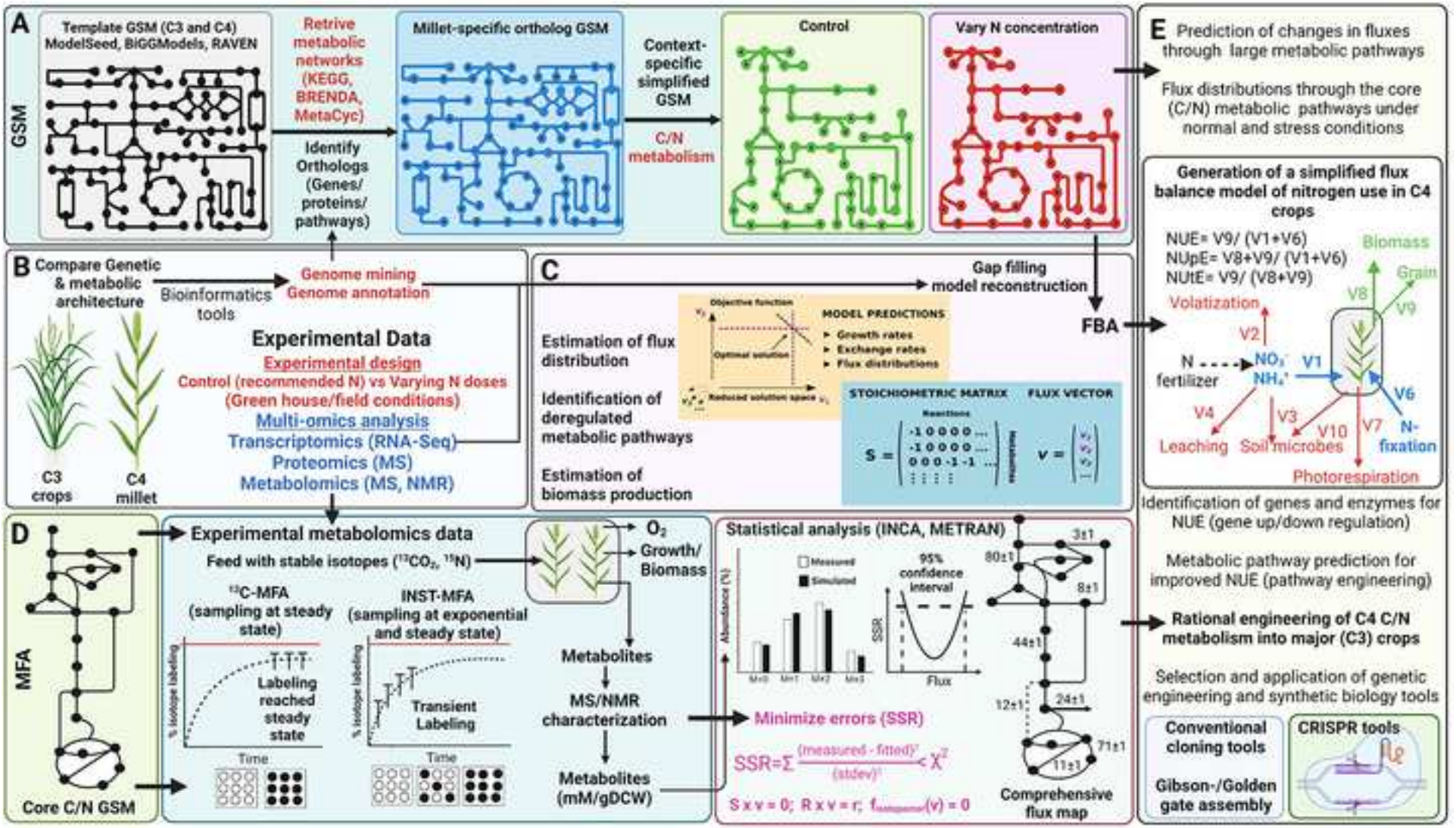
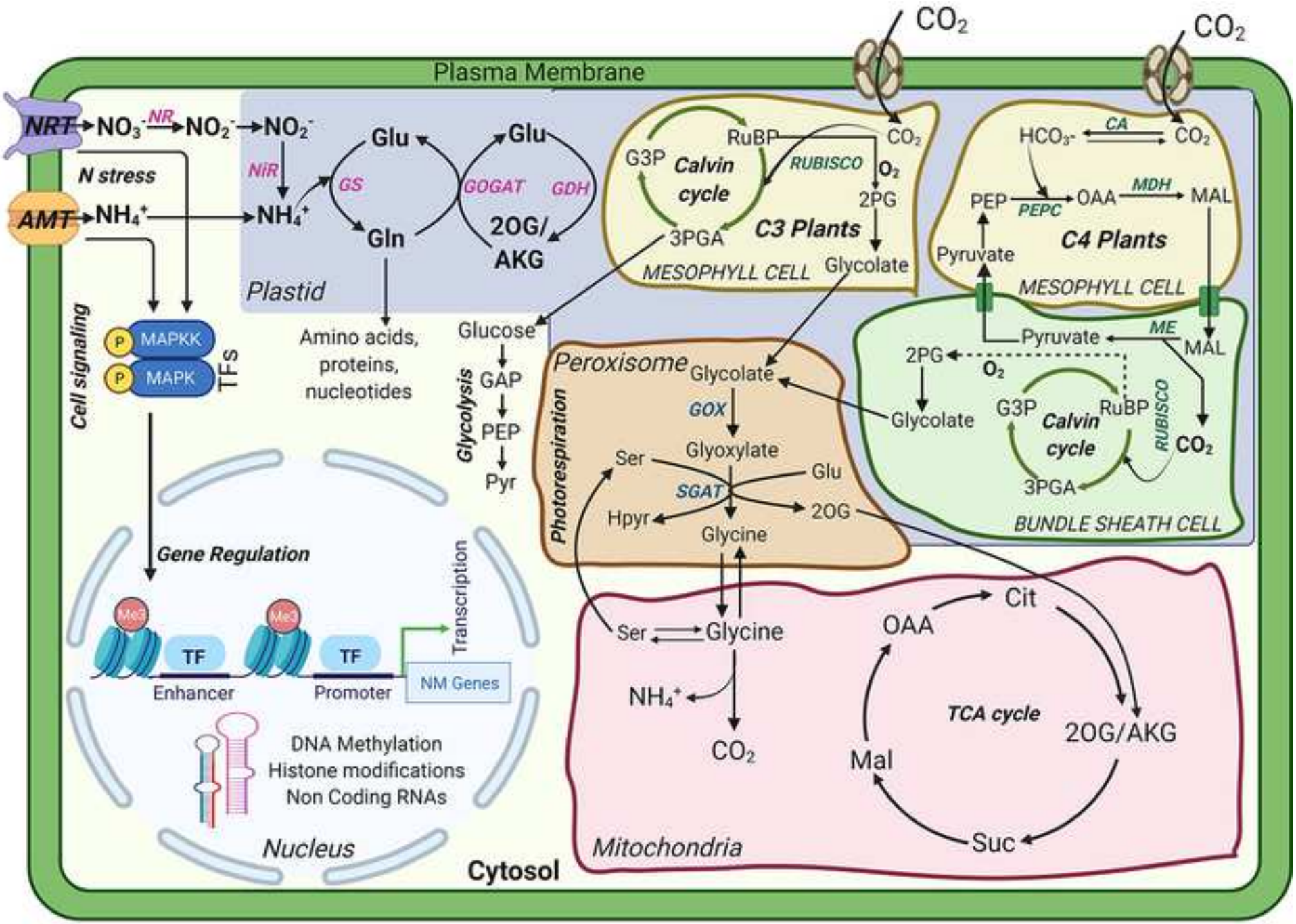


Figure2

Click here to access/download;Figure;Figure2.jpg



Highlights

1. Systems biology shows potential to unravel the mechanism behind the high NUE of orphan crops.
2. INST-MFA makes systems biology study of photoautotrophs and orphan crops feasible.
3. The performance of metabolically engineered cultivars with high NUE can be maximized by the DBTL cycle.
4. Mutants of PII/PII-like proteins and Dof1 transcription factor are promising candidates for studying the systems biology of NUE.
5. Cyanobacterial PII signaling mechanisms provide promising clues for efficient system metabolic engineering of crops.