



# Introduction: Advances in Plant Omics and Systems Biology

1

Flavia Vischi Winck, Lucca de F. R. Monteiro,  
and Glaucia M. Souza

## Abstract

How the complexity of biological systems can be understood is currently limited by the amount of biological information we have available to be incorporate in the vastitude of possibilities that could represent how a biological organism function. This point of view is, of course, alive under the paradigm that describes a living thing as a whole that could never be interpreted as to the sole understanding of its separated parts.

If we are going to achieve the knowledge to understand all the complex relations between the molecules, pathways, organelles, cells, organs, phenotypes, and environments is unknown. However, that is exactly what moves us toward digging the most profound

nature of relationships present in the living organisms.

During the last 20 years, a big workforce was dedicated to the development of techniques, instruments, and scientific approaches that guided a whole new generation of scientists into the universe of omics approaches. The implementation of technological advances in several omics applications, such as transcriptomics, proteomics, and metabolomics, has brought to light the information that nowadays reshape our previous thinking on specific aspects of plant sciences, including growth, development, organ communication, chromatin states, and metabolism, not to mention the underpinning role of regulatory mechanisms that in many cases are essentially the basis for the phenotypical expression of a biological phenomenon and plants adaptation to their environment.

In this chapter, some of the original concepts of complex systems theory were briefly discussed, and examples of omics approaches that are contributing to uncovering emergent characteristics of plants are presented and discussed. The combination of several experimental and computational or mathematical approaches indicated that there is room for improvements and novel discoveries. However, the level of complexity of biological systems seems to require and demand us to unify efforts toward the integration of the large omics datasets already available and the

F. V. Winck (✉)

Institute of Chemistry, University of São Paulo,  
São Paulo, Brazil

Center for Nuclear Energy in Agriculture,  
University of São Paulo, Piracicaba, Brazil  
e-mail: [winck@cena.usp.br](mailto:winck@cena.usp.br)

Lucca de F. R. Monteiro  
Institute of Chemistry, University of São Paulo,  
São Paulo, Brazil

Institute of Biosciences, University of São Paulo,  
São Paulo, Brazil

G. M. Souza  
Institute of Chemistry, University of São Paulo,  
São Paulo, Brazil

ones to come. This unification may represent the necessary breakthrough to the achievement of the understanding of complex phenomena. To do so, the inclusion of systems biology thinking into the training of undergraduate and graduate students of plant sciences and related areas seems to be also a contribution that is necessary to be organized and implemented in a worldwide scale.

### Keywords

Single cell · Networks · Integration · Regulation · Metabolism · Signaling

## 1.1 Overview of Systems Theory Applied to Plant Sciences

The extraordinary complexity of cellular responses of any living organism has long been an attractive topic to scientists, including the understanding of what complexity is and how it is structured in life.

In the endless endeavor of investigating all factors that act concurrently in the definition of cell fate, several new discoveries were made. In special, at the beginning of the twentieth century, several initiatives reemerged to foment the integrative (or holistic) investigation of living organisms, with a special dedication to transdisciplinary studies of animals and plants (Drack et al. 2007) that recapitulated the different views on system theory and thinking and a new modern time emerged in system theories applied to biology. The basic concepts that permeated the system thinking resided in the premise that the properties of the whole cannot be completely understood from the simple sum of the properties of its isolated components (Von Bertalanffy 1972).

Therefore, the systems property should emerge from the enormous number of dynamic interactions (direct and/or indirect) and relations between the components (e.g., molecules, organelles, organs, tissues) of the living organisms, which indeed represent the complexity of the biological systems and that also permits it to

dynamically respond to environmental changes and internal perturbations, which is one of the most important characteristics of the living organisms.

One of the special messages that remained from the studies of the previous century and conceptual elements generated thereof is the need for transdisciplinary thinking and methodologies that could guide us into the understanding of complex phenomena.

In plant sciences, several efforts have been made to address complex phenomena, bringing novel insights into the structure, organization, regulation, and evolution of plant responses. The typical research on systems biology usually proposes the cyclic implementation of experimental and theoretical analysis of the biological systems with the investigation of dynamic biological responses. Experiments are performed in laboratories and computational and/or mathematical modeling and simulation of the biological systems are applied in an iterative manner (Kitano 2002). This approach will certainly enhance our knowledge of complex systems by gradual increments and additional breakthroughs that improve our understanding of underlying principles of the whole biological systems.

The advancement of molecular biology approaches and technologies, including the evolution of different types of OMICS analysis opened new venues for the analysis of the several layers of biological information (here referring to pools of different types of molecules) toward the reconstruction, generation, and validation of new proposed models of biological organisms. The progress and dissemination of OMIC approaches and high-throughput methods are intimately connected to the advances in plant systems biology (Provar and McCourt 2004), and several important initiatives have provided significant contributions of novel biological data, construction, and maintenance of biological databases that could be confidently applied in systems biology approaches (Falter-Braun et al. 2019). The underlying cellular mechanisms that define plant phenotypes are now being investigated through a systems approach in different plant models by researchers from several countries.

## 1.2 Advances in Omics and Systems Biology Applications in Plant Sciences

More than 10 years ago, proposed modeling approaches that integrated chemical and mechanical information have revealed important aspects of plant development, such as meristem development into shoots or flowers, and based on molecular modeling, methodologies have also revealed the challenging aspects of such studies, such as the need of time-resolved information and imaging of biological phenomena integration between molecular data into mechanical, phenotypic descriptions of plant growth (Chickarmane et al. 2010).

Essential processes such as chromatin remodeling and regulation of activation or deactivation of transcriptional units or modules were investigated through genome-wide analysis and meta-analysis of several datasets of *Arabidopsis* plants histone modifications. The analysis performed through a system view brought to light the high complexity of chromatin remodeling processes, evidencing the existence of nine possible different chromatin states. Computational analysis of biological data originated from published profiles of histone modifications, histone variants, nucleosome density, genomic G + C content, CG methylated residues, and chromatin immunoprecipitation (ChIP) data for histone acetylation was performed in an integrative fashion and revealed transcriptional active sites, repressed sites, elongation signatures, intergenic upstream promoter regions, Polycomb, intragenic regions associated to short and long transcript units and two heterochromatin profiles related to intergenic regions and pericentromeric regions. This discovery also revealed the correlation between these regions and gene expression activation or deactivation, which rendered knowledge with a higher confidence of topological organization of chromatin regions and their association or close proximity to each other (Sequeira-Mendes et al. 2014). This type of information thus suggests that chromatin remodeling processes may have a

mechanism or structured biological information that influence a priori the position of the main epigenetic modifications, defining the chromatin topology in the plant cells.

These findings are intrinsically connected with the understanding of the multi-combinatorial nature of the control of gene expression. In *Arabidopsis*, for instance, it has been evidenced that most gene promoters (63%) are recognized by at least two Transcription Factor (TFs) proteins, while some promoters may be recognized by up to 18 different TFs, composing a highly interconnected hub of molecular interactions. Genes that are expressed in many different conditions are usually controlled by many different TFs (Brkljacic and Grotewold 2017).

The integration of spatial information of chromatin signatures and topology with the most recent findings of the distribution of cis-regulatory modules (CRMs) along the genomes may substantially reveal some principles of the global regulation of gene expression and causative structures that can be associated with phenotypes of interest, such as plant growth and stress responses. An exciting review on gene expression control can be consulted in the work from Brkljacic and Grotewold (2017). Some interesting questions arise from such studies on gene combinatorial expression, including the possible global and conserved preferences of groups or families of TFs for binding to correlated chromatin regions and states, the effects of TFs in the regulation of non-available CRMs and how the different groups of TFs are associated to work together in multiple different complexes, depending on the CRMs and chromatin they interact.

In this present book, you will find the presentation of interesting examples of epigenetic mechanisms and its principles, the basics of interaction networks and the strategies to generate models that can be implemented to describe the possible connections of the elements of the biological systems.

In a different but complementary perspective, the analysis of the transcripts and proteins expressed in a cell or tissue will be introduced and advances discussed. These omics data have rendered a massive amount of qualitative and

quantitative biological information in the last two decades. The omics approaches revealed, in several instances, with a time resolution, the dynamics of the cellular responses with the indication of timely coordinated events, cyclic, inhibited, and induced responses of the gene complements (transcripts and proteins).

As a side note, the growing volume of multi-omics information, together with higher computational processing capabilities, states a duality that must be addressed: the enhancement of the integration of different layers of biological information from different datasets towards the improvement of the understanding of complex systems. This requires extensive communication, the development of databases and data sharing between researchers. The organization of multinational participative computational repositories and open access data analysis platforms guided to continuous data mining and data integrative analysis improvement could integrate the plant systems biology community around the problems that concern the major tasks of interpreting the complex systems. This would add a social benefit of contributing to expanding the access of non-developed countries to the advanced science in the field of plant biology and computational biology.

In such context, the *Crops in silico* initiative (<https://cropsinsilico.org/>) represents an interesting prospect for integrative modeling tools in plant sciences. In fact, there is a shortage of plant-based multi-scale simulations compared to the number of models of mammals, with the existing plant models being restricted to time-limited descriptions of several singular biological processes or phenotypic responses to environmental stimuli. There is indeed an urge for the development of a virtual physiological plant, such a model also integrating developmental timescales and environmental data to plant multi-omic networks and phenotypes to understand response complexities. The *initiative* tries to address such demands through constructing a plant community-centered platform, while also dealing with usual collective technical barriers: visualization, data imputation, coding standardizations, and accessibility issues (Marshall-Colon et al. 2017).

The massive scale of current transcriptome data analysis, in special for model plants such as *Arabidopsis thaliana*, resulted in the identification of many different molecular phenotypes, generating novel insights on how changes in the transcriptional state of the cells are associated with global patterns of gene expression control. For instance, time-resolved transcriptome analysis of the *Arabidopsis* root revealed that different nitrogen doses induced the modulation of 1153 genes in a pattern that fits a Michaelis–Menten kinetics, indicating the existence of a saturation trend of transcripts accumulation or depletion at upper levels of nutrient availability. This study also revealed that some early responsive TF genes are likely related to compound-dose-responsive transcription (Swift et al. 2020). Even though the existence of a gradual transcriptional response has been proposed for a while, the deep investigation of this type of global patterns is necessary, since the promoter architecture of several genes revealed the presence of repeated binding sites which suggested their capacity to interact with several protein complexes containing TFs (Brady et al. 2006).

These results raised some questions about how much of the gene expression control can directly be affected by the concentration of external compounds following a simple kinetics and if this type of control is happening through a direct or indirect way at the DNA level in different environmental conditions. In the same way, it is essential to know how these global patterns are established and conserved in the cells, in a multi-combinatorial regulatory network structure, where multiple TFs may bind to multiple *cis*-regulatory elements. The understanding of these mechanisms can contribute to the identification of the nature of the coordination of these modular transcriptional states and their integration into new cellular functions, opening a new path for identification and modeling of causative effects on cell functional regulation. In addition, the several transcriptome analyses performed in the past 15 years have also contributed to reveal several molecular phenotypes in detail that exposed the connection of many biological processes. For instance, the transcriptome depicted genome-

wide oscillations of gene expression, with the identification of genes essential for circadian rhythm and photosynthesis, cell growth, and division, contributed to the identification of the promoter elements correlated to the genes unexpectedly modulated under circadian control (Harmer et al. 2000). More recently, the involvement of pre-mRNA processing, transcript stability, mRNA nuclear export, posttranslation, and non-protein coding RNAs (ncRNAs) in particular, long ncRNAs (lncRNAs), in the regulation of circadian rhythm in plants (Romanowski and Yanovsky 2015) exposed the complexity of this biological process compared to the first model of transcriptional–translational feedback loop described in plants with the participation of two *Arabidopsis* MYB transcription factors CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) and LATE ELONGATED HYPOCOTYL (LHY) (Alabadi et al. 2001; Schaffer et al. 1998; Strayer et al. 2000; Wang and Tobin 1998).

In the light of evolution, the recent efforts on the analysis of more than one thousand plant transcriptomes opened a new path to a comprehensive understanding of evolutionary differences between plants and their cellular mechanisms. The integration of data on plant habitat and niches with molecular data may reveal novel sets of genes involved in specific adaptations and phenotypes. The broad phylogenetic analysis based on transcriptome data also indicated that most gene expansion events in plant lineage have occurred before the appearance of vascular plants (One Thousand Plant Transcriptomes 2019), intriguing by its great possibilities of future applications in plant biotechnology.

In addition, computational resources implemented for plant functional genomics data visualization and mining such as Bio-analytic Resource (BAR) for Plant Biology with eFP Browser (Winter et al. 2007) have been of great benefit to the plant community to interpret such vastitude of data. The same is true for other platforms such as Phytozome (Goodstein et al. 2012) that over the years have made available hundreds of genomic datasets and have expanded its applications into Phytomine ([https://phytozome.jgi.](https://phytozome.jgi.doe.gov/phytomine/begin.do)

[doe.gov/phytomine/begin.do](https://phytozome.jgi.doe.gov/phytomine/begin.do)), integrating genomics and functional data and fostering data mining. Gramene is also an example bringing integration to the plant reactome (<http://plantreactome.gramene.org/index.php?lang=en>). The current expansion of such platforms to integrate other types of omics data would be beneficial to the future of plant systems biology applications.

Proteomics approaches have also revealed important aspects of plant phenotypes, especially the molecular description of metabolic pathways operating in cellular responses, and significantly increased the number of proteins identified that are related to crops productivity and stress responses (Salekdeh and Komatsu 2007). The last 15 years have witnessed an expansion of proteomics data ranging from land plants to unicellular algae, showing the particularities of these organisms in responses to variations of environmental conditions such as light (Mettler et al. 2014), CO<sub>2</sub> content (Santos and Balbuena 2017), metabolic regimes (Vidotti et al. 2020), among others. It is also noticeable that proteomics is opening venues for a broader understanding of the cell response regulation by revealing the identity and possible regulatory role of many proteins that undergo posttranslational modifications (Huang et al. 2019; Van Leene et al. 2019). The quantitative proteomic profiling of cellular responses has been applied to model plant species, and it is quickly expanding to cover non-model species, which is highly desirable and necessary to uncover the vast diversity of metabolic characteristics and nuances of the plant species naturally adapted to different environmental niches, such as desertic, tropical, semiarid, and rainy regions. This and other topics of plant proteomics are presented and discussed within this book.

The transferable application of omics knowledge into system-biology-based plant breeding is one important consequence of the development of large biological datasets of omics and plant systems biology data (Lavarenne et al. 2018). Natural breeding or genetic engineering of plants should now address great problems with a more holistic approach of plant systems contributing to the generation of novel stress-resistance crops (Zhang et al. 2018).

In a parallel trend, similar paradigmatic breakthroughs in plant systems biology have been recently achieved through metabolomics studies. Together with improved mass spectrometry (MS) techniques, advancements in computational biology and bioinformatics have been allowing for wider modeling efforts, permitting broader integration of metabolite data with other large amounts of transcriptome and proteome information, further bolstering the construction of interactomes. This has important consequences given the massive size of the plant metabolome and the fact that it is usually the first layer of biological information within the cell to be subjected to the effects of environmental changes since the response at transcriptional level can take longer to occur, if not to mention the direct influence of metabolites in transcriptional modulation and other possible interactions of the metabolome with transcripts and proteins.

An applied sample of this trend can be noted through the work of Veyel and collaborators with a study in *Arabidopsis* where they developed an improved proteomic-compatible, metabolomic-oriented method for system-wide analyses of protein-small molecule complexes. The protein-metabolite interactome is an often overlooked but functionally important regulatory feature to be dealt with when studying metabolite-diverse, metabolome data-rich systems such as plants—retroactively, this substantial amount of information also makes data collecting and processing troublesome. A simple, but innovative addressing to this issue is the proposed general-case, large-scale analysis-oriented co-fractionation method instead of canonical approaches such as cross linking or protein tagging for search of interaction sites. The approach relies on the hypothesis that proteins and protein-bound small molecules fractionate together when forming stable complexes: as such, the use of size separation techniques should concentrate protein-metabolite complexes in higher molecular weight fractions, the process could then be coupled with analytic techniques (i.e., mass spectrometry). Apart from the technological prospect, this proof-of-concept has also identified a plethora of novel stable protein-metabolite complexes from the

*Arabidopsis* samples, suggesting emergent regulatory roles for some small molecules, with potential for extension to other biological systems (Veyel et al. 2017).

The further study of how the metabolic landscape of plants is shaped by varying environmental conditions can also unravel the metalinguistics of experimental design in plant sciences. A common consequence of the application of high-throughput analytical techniques to biological systems is the finding that often formerly overlooked properties of the environment can exert unexpected effects in an organism's metabolism, potentially biasing the reproducibility of some experiments. As a matter of fact, sensitivity to initial conditions is an inherent condition of complex systems, particularly in a multi-omics perspective.

While dealing with experimentation in plants, lighting conditions represent an essential, though often taken for granted environmental condition. Sunlight is characterized by sinusoidal changes in irradiance throughout the day cycle, with shading and clouds momentarily varying the amount of light absorbed by the plant. On the other hand, artificially lighted growth chambers usually offer constant light irradiance (square wave), abrupt light–dark shifts and different spectral quality when compared to naturally lighted environments. Ironically, although growth chambers are considered essential for experimental reproducibility for scientific approaches, the vastitude of differences from the phenotypes of plants grown in such environments with the phenotypes observed in plants in natural environments and in the field enhance the number of possible dynamic phenotypes that may populate the universe of metabolomics.

This issue is illustrated through metabolite analysis of samples obtained from *Arabidopsis* plants grown under greenhouse and growth chamber (with sinusoidal or square lighting patterns, fluorescent or sunlight spectrasimulating LED light) conditions. The combination of enzymatic assays with HPLC and LC-MS indicates major differences (fold change) in occurrence of components of central carbon and nitrogen pathways when data from

greenhouse and growth chamber conditions are compared altogether. In the context of adaptation difficulties faced by the plant while dealing with unexpected changes in sunlight within and between days, emphasis should be placed on the importance of variations (for greenhouse versus growth chamber experiments) in photo-assimilate partitioning optimization, amino acid synthesis (Ser, Gln, Gly), C:N ratio, and synthesis of components of sugar signaling networks such as sucrose and the sucrose status indicator, trehalose 6-phosphate, those finely linked with the sugar feedback-operated regulation of starch metabolism (turnover) and sucrose homeostasis through day–night cycles by the circadian clock in *Arabidopsis*: the circadian oscillator itself suffering adjustments through feedback from sunlight pattern changes (Seki et al. 2017). The metabolite analysis also reveals that milder differences were observed when comparing different conditions of artificial illumination (i.e., sinusoidal versus square patterns) and that LED lighting may not fully represent natural lighting conditions (Annunziata et al. 2017).

In this present book, different aspects of metabolomics basics and applications are introduced, discussing scientific and methodological approaches that are contributing to broadening the knowledge on the metabolic regulation of plant phenotypes. Nevertheless, the metabolomics analysis still faces a great challenge of identifying the cellular compartmentalization of the metabolites in the different cell responses, which may include an extraordinary level of complexity in the cellular responses, especially if other dual interactions of metabolites, such as metabolite–protein or metabolite–miRNAs, occur in a dimension of hundreds of thousands.

---

### 1.3 Challenges in Plant Systems Biology and Paths to Expand the Research Field

Besides their sensitivity to initial conditions, biological systems are also characterized by emergence. The interaction between parts of a system

can generate emergent properties, sometimes loosely linked with such parts. As a remarkable example, the use of mass spectrometry (affinity purification-mass spectrometry, gas chromatography-mass spectrometry) has revealed that glycolytic enzymes can mediate mitochondria-chloroplast colocalization in *Arabidopsis*. This finding sums to an already extensive array of supposed properties of the glycolytic pathway in non-plant cells, such as the formation of multienzyme complexes, colocalization with ATP-demanding areas and enzyme chemotactic movements, not to mention the physical association of glycolytic enzymes with the mitochondria (Zhang et al. 2020). Along stating that even well-studied central pathways can perform unprecedented roles, one can also argue that the given groundbreaking work—in a similar fashion to other featured research works—asserts the insufficiency of reductionism to deal with complexity. This context has a special impact on undergraduate biochemistry disciplines.

The known biological complexity substantially impacts experimental design in plant molecular physiology studies. In cultivated crops where multiple copies of genomes are present, the computational challenges alongside those of mathematics and statistical genetics are indeed formidable. The investigation of emergent properties emphasizes to us the importance of the chosen approach to teach complexity in some life science courses. Despite its non-intuitive nature, emergence is an important component of the Dynamical Systems Theory, which constitutes one of the three systems theories, along with Bertalanffy's General Systems Theory and Cybernetics: the so-called systems thinking concept has regained recent relevance in primary and secondary education. Its crosscutting characteristic and multidisciplinary applicability are especially helpful in teaching skills to comprehend biological complexity (Verhoeff et al. 2018). Yet, there is comparatively less extensive research and application of systems thinking in STEM education, a substantial lack of integration with chemistry courses is particularly noticeable while most of existing peer-reviewed literature on systems thinking generally focuses on biology education.

The implementation of systems thinking approach in areas such as chemistry and biochemistry may enhance student's strength in taking complex decisions that are important for global issues, such as sustainability (York et al. 2019; York and Orgill 2020). Although not extensively, some life science courses have recently been incorporating systems theory concepts into their programs, a trend exemplified by attempts to apply ecological perspectives and systems modeling into redesigning an undergraduate botany major course (Zangori and Koontz 2017). This is specially targeted at meeting novel demands on teaching biological complexity and retroactively generates content-changing demands for the disciplines that service those courses: still, while paramount to most life science careers, biochemistry is such a complicated example.

The inclusion of space for the teaching of the basics of systems biology-related topics such as complex system theories that could be explored through the concepts of the regulatory landscape of biological systems (i.e., molecular binding to sites in proteins and nucleic acid sequences as a topic for structure–function classes, transcriptional regulation, posttranslational modifications), multi-omics networks, high-throughput techniques for data acquisition (mass spectrometry, sequencing), and computational analysis would contribute to change the oftentimes discrete, linear fashion for teaching the structure, components and behavior of certain well-studied, canonical cellular mechanisms and pathways. Although time and resource constraints indeed interfere in which contents (and to what extent they can be deepened) should be taught, one could argue that an aged linear, pathway-focused attempt to teach basic biochemistry stretches away from thoroughly portraying complexity by snatching its integrative, network-based, and emergent attributes.

Systems biology and omics approaches can also be wisely applied to explore more mundane, unforeseen problems in teaching scenarios. The implementation of basic bioinformatic skills in early years of the undergraduate courses in bio-

logical and biochemistry sciences is essential to prepare the students for the future of populated biological databases and data mining schemes that are going to permeate their future academic and professional lives. Examples on how to connect the systems biology and omics knowledge into the curriculum and lectures for undergraduate students can cross the successful examples of omics applications in the several fields of research, including plant sciences and the more recent synthetic biology approaches for bioproducts production. Therefore, we envisage that some examples of studies presented and discussed in the present book may in the future be applied in systems biology-based graduate and undergraduate schools worldwide.

---

## References

- Alabadi D et al (2001) Reciprocal regulation between TOC1 and LHY/CCA1 within the Arabidopsis circadian clock. *Science* 293(5531):880–883
- Annunziata MG et al (2017) Getting back to nature: a reality check for experiments in controlled environments. *J Exp Bot* 68(16):4463–4477
- Brady SM, Long TA, Benfey PN (2006) Unraveling the dynamic transcriptome. *Plant Cell* 18(9):2101–2111
- Brkljacic J, Grotewold E (2017) Combinatorial control of plant gene expression. *Biochim Biophys Acta Gene Regul Mech* 1860(1):31–40
- Chickarmane V et al (2010) Computational morphodynamics: a modeling framework to understand plant growth. *Annu Rev Plant Biol* 61:65–87
- Drack M, Apfalter W, Pouvreau D (2007) On the making of a system theory of life: Paul A Weiss and Ludwig von Bertalanffy's conceptual connection. *Q Rev Biol* 82(4):349–373
- Falter-Braun P et al (2019) iPlant Systems Biology (iPSB): an international network hub in the plant community. *Mol Plant* 12(6):727–730
- Goodstein DM et al (2012) Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res* 40(Database issue):D1178–D1186
- Harmer SL et al (2000) Orchestrated transcription of key pathways in Arabidopsis by the circadian clock. *Science* 290(5499):2110–2113
- Huang D et al (2019) Protein S-nitrosylation in programmed cell death in plants. *Cell Mol Life Sci* 76(10):1877–1887
- Kitano H (2002) Systems biology: a brief overview. *Science* 295(5560):1662–1664
- Lavarenne J et al (2018) The spring of systems biology-driven breeding. *Trends Plant Sci* 23(8):706–720



- Marshall-Colon A et al (2017) Crops in silico: generating virtual crops using an integrative and multi-scale modeling platform. *Front Plant Sci* 8:786
- Mettler T et al (2014) Systems analysis of the response of photosynthesis, metabolism, and growth to an increase in irradiance in the photosynthetic model organism *Chlamydomonas reinhardtii*. *Plant Cell* 26(6):2310–2350
- One Thousand Plant Transcriptomes, I (2019) One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* 574(7780):679–685
- Provart NJ, McCourt P (2004) Systems approaches to understanding cell signaling and gene regulation. *Curr Opin Plant Biol* 7(5):605–609
- Romanowski A, Yanovsky MJ (2015) Circadian rhythms and post-transcriptional regulation in higher plants. *Front Plant Sci* 6:437
- Salekdeh GH, Komatsu S (2007) Crop proteomics: aim at sustainable agriculture of tomorrow. *Proteomics* 7(16):2976–2996
- Santos BM, Balbuena TS (2017) Carbon assimilation in *Eucalyptus urophylla* grown under high atmospheric CO<sub>2</sub> concentrations: a proteomics perspective. *J Proteome* 150:252–257
- Schaffer R et al (1998) The late elongated hypocotyl mutation of *Arabidopsis* disrupts circadian rhythms and the photoperiodic control of flowering. *Cell* 93(7):1219–1229
- Seki M et al (2017) Adjustment of the *Arabidopsis* circadian oscillator by sugar signalling dictates the regulation of starch metabolism. *Sci Rep* 7(1):8305
- Sequeira-Mendes J et al (2014) The functional topography of the *Arabidopsis* genome is organized in a reduced number of linear motifs of chromatin states. *Plant Cell* 26(6):2351–2366
- Strayer C et al (2000) Cloning of the *Arabidopsis* clock gene *TOC1*, an autoregulatory response regulator homolog. *Science* 289(5480):768–771
- Swift J et al (2020) Nutrient dose-responsive transcriptome changes driven by Michaelis-Menten kinetics underlie plant growth rates. *Proc Natl Acad Sci U S A* 117(23):12531–12540
- Van Leene J et al (2019) Capturing the phosphorylation and protein interaction landscape of the plant TOR kinase. *Nat Plants* 5(3):316–327
- Verhoeff RP et al (2018) The theoretical nature of systems thinking. *Perspectives on systems thinking in biology education*. *Front Educ* 3:40
- Veyel D et al (2017) System-wide detection of protein-small molecule complexes suggests extensive metabolite regulation in plants. *Sci Rep* 7:42387
- Vidotti ADS et al (2020) Analysis of autotrophic, mixotrophic and heterotrophic phenotypes in the microalgae *Chlorella vulgaris* using time-resolved proteomics and transcriptomics approaches. *Algal Res* 51:102060
- Von Bertalanffy L (1972) The history and status of general systems theory. *Acad Manag J* 15(4):407–426
- Wang Z-Y, Tobin EM (1998) Constitutive expression of the *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*) gene disrupts circadian rhythms and suppresses its own expression. *Cell* 93(7):1207–1217
- Winter D et al (2007) An “Electronic Fluorescent Pictograph” browser for exploring and analyzing large-scale biological data sets. *PLoS One* 2(8):e718
- York S, Orgill M (2020) ChEMIST Table: a tool for designing or modifying instruction for a systems thinking approach in chemistry education. *J Chem Educ* 97(8):2114–2129
- York S et al (2019) Applications of systems thinking in STEM education. *J Chem Educ* 96(12):2742–2751
- Zangori L, Koontz JA (2017) Supporting upper-level undergraduate students in building a systems perspective in a botany course. *J Biol Educ* 51(4):399–411
- Zhang H, Li Y, Zhu JK (2018) Developing naturally stress-resistant crops for a sustainable agriculture. *Nat Plants* 4(12):989–996
- Zhang Y et al (2020) A moonlighting role for enzymes of glycolysis in the co-localization of mitochondria and chloroplasts. *Nat Commun* 11(1):4509