






INVITED REVIEW

Systems biology for crop improvement

Lekha T. Pazhamala^{1,*}  | Himabindu Kudapa^{1,*}  | Wolfram Weckwerth^{2,3}  |
 A. Harvey Millar⁴  | Rajeev K. Varshney^{1,5} 

¹ Center of Excellence in Genomics & Systems Biology, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, Hyderabad 502 324, India

² Department of Ecogenomics and Systems Biology, University of Vienna, Vienna, Austria

³ Vienna Metabolomics Center, University of Vienna, Vienna, Austria

⁴ ARC Centre of Excellence in Plant Energy Biology and School of Molecular Sciences, The University of Western Australia, Perth, WA, Australia

⁵ State Agricultural Biotechnology Centre, Crop Research Innovation Centre, Food Futures Institute, Murdoch University, Murdoch, WA, Australia

Correspondence

Rajeev K Varshney, Center of Excellence in Genomics & Systems Biology, ICRISAT, Patancheru- 502 324, Hyderabad, India
 Email: r.k.varshney@cgiar.org,
rajeev.varshney@murdoch.edu.au

*Both authors contributed equally: Lekha T Pazhamala and Himabindu Kudapa

Assigned to Associate Editor Henry T. Nguyen.

Funding information

Science & Engineering Research Board (SERB) of Department of Science & Technology (DST), Government of India, Grant/Award Numbers: SB/S9/Z-13/2019, SB/WEA-01/2017; Bill and Melinda Gates Foundation, Grant/Award Number: OPP1130244

Abstract

In recent years, generation of large-scale data from genome, transcriptome, proteome, metabolome, epigenome, and others, has become routine in several plant species. Most of these datasets in different crop species, however, were studied independently and as a result, full insight could not be gained on the molecular basis of complex traits and biological networks. A systems biology approach involving integration of multiple omics data, modeling, and prediction of the cellular functions is required to understand the flow of biological information that underlies complex traits. In this context, systems biology with multiomics data integration is crucial and allows a holistic understanding of the dynamic system with the different levels of biological organization interacting with external environment for a phenotypic expression. Here, we present recent progress made in the area of various omics studies—integrative and systems biology approaches with a special focus on application to crop improvement. We have also discussed the challenges and opportunities in multiomics data integration, modeling, and understanding of the biology of complex traits underpinning yield and stress tolerance in major cereals and legumes.

Abbreviations: BPH, brown planthopper; GEA, gene expression atlas; GWAS, genome-wide association study; NGS, next-generation sequencing; QTL, quantitative trait loci; SNP, single nucleotide polymorphism; SV, structural variation.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *The Plant Genome* published by Wiley Periodicals LLC on behalf of Crop Science Society of America

1 | INTRODUCTION

Crops such as cereals and legumes play an important role in human diet by providing necessary calories, proteins, essential amino acids, and minerals. Over the last 100 yr of extensive breeding, crop varieties have been developed for higher yields. However, because of increasing global population, there is an urgent need to double the yield by the year 2050. This increase in production and productivity is challenging considering the current environmental constraints and rapidly depleting natural resources. Model plants, such as *Arabidopsis* (*Arabidopsis thaliana* (L.) Heynh.) and rice (*Oryza sativa* L.), have been extensively studied for comprehensive understanding of plant genetics, genomics, and defining the function of specific genes. This is essential for leveraging genomics for breeding a new generation of climate-ready crops to produce surplus food that is high in nutrients. Toward this, a genomics-assisted breeding (Varshney et al., 2005) approach can greatly accelerate the existing crop improvement programs. Translation or transfer of genetic information gained from one species to another was quite limited before the genomics era, mainly because of a lack of suitable knowledge in genomic information and systems biology.

The advent of next-generation sequencing (NGS) technologies has revolutionized and increased the pace of generation of genomics and transcriptomics data that has led to new era of the ‘big data.’ Several NGS platforms, such as Illumina’s MiSeq/HiSeq; Roche’s 454/FLX; ABI/Life Technologies’ SOLiD; Invitrogen’s Ion Proton, have led to the sequencing of genomes and transcriptomes for a number of plant species (Barutcu et al., 2015; van Dijk et al., 2018). Third-generation sequencing technologies such as single-molecule sequencing by Helicos Biosciences (HeliScope), single molecule real-time sequencing by Pacific Biosciences (PacBio), and Nanopore sequencing by Oxford Nanopore Technologies accelerated generation of large-scale sequencing data (Giani et al., 2020). This has dramatically changed the sequencing scenarios and led to the development of high-quality genome assemblies in several crop plants including complex and large-sized genomes. The big data from omics experiments are analyzed with advanced software programs and analytical methods to understand the complexity of biological systems. This new area of research has come to be known as either integrative biology or systems biology. Integrative biology focuses on combining omics layers to build insight based on the Aristotelian principle that the whole is greater than the sum of the parts. Systems biology focuses on combining omics layers to build models that explain system behavior and that have predictive power to propose the outcome of mutation or modification of specific biological steps. Furthermore, systems biology has applications in dissecting complex agronomic traits and in the model-based prediction of phenotypes in different conditions (Lavrenne

Core Ideas

- Multiomics data enable understanding and prediction of plant phenotype with higher precision.
- Systems biology approach involves multiomics data integration and modeling.
- Integrative systems biology aids better prediction of cellular functions of complex traits.
- Advancements, challenges, and opportunities in systems biology research are reviewed.
- Prospects of systems biology to crop improvement are presented.

et al., 2018). The development of user-friendly pipelines and bioinformatic tools to analyze the big data generated by omics approaches can further facilitate breeding programs both through enhanced selection tools and through more sophisticated design of crossing programs or stacking of gene modifications.

An overview of current omics resources, multiomics data integration, and systems biology approaches each with a focus on their applications in plant research and breeding has been discussed in the present review (Figure 1). We highlight existing and emerging approaches that contribute to our understanding of the biology of complex traits and holistic improvement of yield together with tolerance and resistance to abiotic and biotic stresses. Furthermore, the prospects and challenges facing multiomics data integration, modeling, and systems-level analyses, particularly with the fast-emerging omics technologies, have been discussed. The thoughts presented in this review provide insights on applications of integrated multiple omics research and systems biology for crop improvement.

2 | RAPIDLY EVOLVING OMICS APPROACHES

High-throughput technologies have revolutionized plant research through the study of a whole set of biological entities, including DNA, RNA, proteins, metabolites, and others, of a given species and have been noteworthy. This high-throughput measurement has led to an array of approaches carrying the omics suffix such as genomics, pangenomics, transcriptomics, proteomics, metabolomics, epigenomics, and, more recently, single-cell omics, phenomics, and QTL-omics. These approaches are now integrated across multiple omics layers, providing an opportunity to understand the flow of information that underlies trait biology. In the following sections, we will be discussing these rapidly evolving omics approaches for crop improvement.

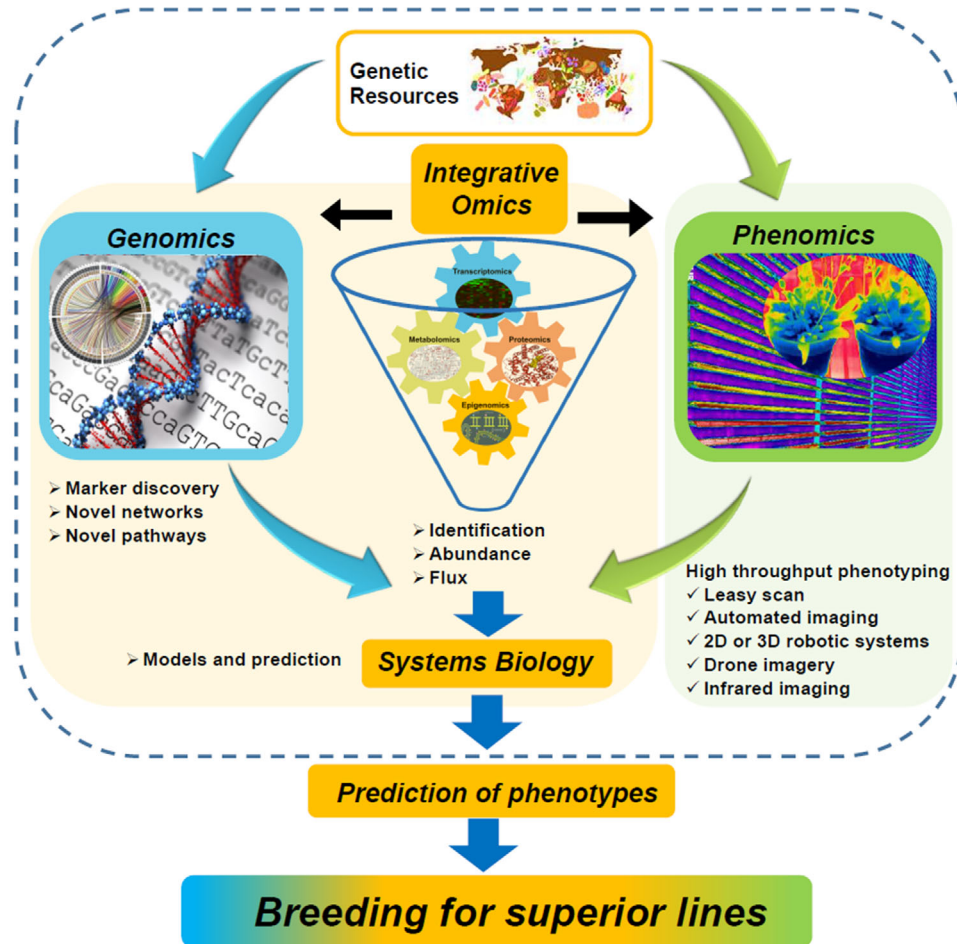


FIGURE 1 An overview of the integrated omics in genetics and breeding for crop improvement. ‘Omics revolution’ is an integrated comprehensive omics approach combining genomics, transcriptomics, proteomics, metabolomics, epigenomics, and other breeding tools for advancement of systemic sciences and to accelerate genomics-enabled, next-generation breeding

2.1 | Genomics and pangenomics

2.1.1 | Genome sequencing and resequencing

Understanding the structure, organization, and dynamics of genomes in plant species can provide insights into how genes have been adapted or altered by natural and artificial selection in response to environmental constraints. Studies have demonstrated the potential use of the adapted and manipulated genes for crop improvement not only within a species but also the use of such genes across species (Kawashima et al., 2016). Toward accomplishing these goals, many plant genomes have been sequenced (Kersey, 2019). Earlier, only a few plant species with relatively compact genomes, and known as ‘models,’ were sequenced to understand genome architecture because of the high cost of sequencing and limited expertise. As a result, genome sequence information for *Arabidopsis thaliana* (L.) Heynh., rice (*Oryza sativa* L.), black cottonwood (*Populus trichophora* Torr. & A. Gray),

grapevine (*Vitis vinifera* L.), and maize (*Zea mays* L.) were the first plant genomes generated. Later, legumes such as soybean [*Glycine max* (L.) Merr.], pigeonpea [*Cajanus cajan* (L.) Huth], and chickpea (*Cicer arietinum* L.) genomes were sequenced (see Varshney et al., 2015).

Evolution in NGS and third-generation sequencing technologies has resulted in ever-increasing throughput and reduced sequencing costs. As a result, more than 600 complete plant genome assemblies are available in public repositories (Kersey, 2019) and many more are being sequenced (<http://www.onekp.com/>). The genome sequence information generated through high-throughput sequencing of germplasm collection is also enabling the simultaneous discovery and sequencing of thousands of genetic markers across whole genomes (Varshney et al., 2019a). These new sequencing tools are also valuable for the validation and assessment of genetic markers in populations. Further, it has been possible now to identify all the genes in a plant, which would in turn help understand the genetic properties as

well as networks that contribute to develop superior crop varieties.

In addition to information made available from the genome sequence of cultivars, characterization of genetic diversity present in wild crop relatives and landraces conserved in gene banks are a valuable source of novel genes that could enhance yield and resistance to abiotic and biotic stresses. In this context, resequencing efforts of large-scale germplasm collections have become important. One such example is the 3,010 diverse Asian cultivated rice genomes from the 3,000 Rice Genomes Project (Wang et al., 2018). The study identified 29 million single nucleotide polymorphisms (SNPs), 2.4 million small indels, and over 90,000 structural variations (SVs) that contribute to within- and between-population variation. This study highlighted the genetic diversity that exists in rice germplasm repositories with agriculturally relevant phenotypes. Further the study demonstrated the use of identified SNPs in trait mapping analysis for the highly heritable traits, grain length, grain width, and bacterial blight resistance in rice (Wang et al., 2018). Another example of maize reported sequencing of 278 maize inbred lines and demonstrated extensive variation in SNPs (27 million), indels (287,504), and copy number variations that can potentially be used as a selection index in future maize breeding programs (Jiao et al., 2012). The study showed that modern breeding has introduced dynamic genetic changes into the maize genome, and further artificial selection has affected thousands of targets including genes and nongenetic regions leading to a reduction in nucleotide diversity and an increase in the proportion of rare alleles (Jiao et al., 2012). In the case of dryland cereals, for example, resequencing of 44 sorghum [*Sorghum bicolor* (L.) Moench] lines representing the primary gene pool and spanning dimensions of geographic origin, end use, and taxonomic group resulted in identification of 8 million SNPs, 1.9 million indels, and specific gene loss and gain events for use in sorghum improvement (Mace et al., 2013). This study on sorghum presented that a large untapped pool of diversity exists not only in races of sorghum but also in the allopatric Asian species *S. propinquum* (Kunth) Hitchc. A strong racial structure and complex domestication events were observed with in the accessions studied. Similarly, in pearl millet [*Cenchrus americanus* (L.) Morrone], resequencing of 994 lines resulted in identification of more than 29 million SNPs and 3 million indels for better understanding of trait variation and accelerating genetic improvement (Varshney et al., 2017b). This study highlighted the application of resequencing data to understand the population structure, genetic diversity, and domestication of pearl millet. Further genomic prediction was employed to predict pearl millet hybrid performance and genome-wide association study (GWAS) predicted yield-associated traits in both irrigated and drought conditions.

In legumes, soybean for instance, resequencing of 302 wild and cultivated soybean accessions identified 9 million SNPs and 876,799 indels, providing genes related to domestication and resources for genomics-enabled crop improvement (Zhou et al., 2015). In a different study, 17 wild and 14 cultivated soybean genomes have been resequenced (Lam et al., 2010), thus revealing patterns of genetic variation between wild and cultivated soybeans. This study identified greater allelic diversity in wild soybean and a set of 205,614 SNPs for use in quantitative trait loci (QTL) mapping and association studies. Very recently, Valliyodan et al. (2021) analyzed genetic diversity and structure from the resequencing of 481 diverse soybean accessions, comprising 52 wild selections and 429 cultivated varieties (landraces and elites). This study identified evidence of distinct, mostly independent selection of lineages by particular geographic location. Recently, we have also undertaken the sequencing and phenotyping of thousands of global composite collection of chickpea genomes as part of the 3,000 Chickpea Genome Sequencing Initiative. Under this initiative, resequencing of 429 chickpeas sampled from 45 countries identified a map of 4.97 million SNPs and GWAS identified 262 markers and several candidate genes for 13 different traits associated with drought and heat tolerance mechanisms (Varshney et al., 2019b). Similar efforts were carried out in pigeonpea, where resequencing of 292 pigeonpea accessions resulted in identification of 15.1 million SNPs and 2.1 million indels. This study revealed genomic regions associated with domestication and markers linked with key traits such as flowering time control, seed development, and pod dehiscence (Varshney et al., 2017a). In brief, the sequencing and resequencing studies in several crop species demonstrated the use of genomes and SNPs for trait mapping analyses. Such studies are expected to guide and accelerate crop breeding by identifying genetic variation that will be useful in breeding efforts in different crops and future sustainable agriculture.

2.1.2 | Pangenomics

High-throughput resequencing technologies have been employed in several crops with an aim to explore genomic diversity and to uncover molecular basis of important agronomic traits. However, in all these resequencing studies, characterization of the genetic variants depends on high levels of sequence similarity to map the short reads onto the reference genome, which may miss highly polymorphic regions and regions that are not present in the reference genome (Zhou et al., 2015). Therefore, with an objective to capture all possible variations in a given germplasm collection of a particular species, pangenomics studies have been conducted in several species (see Khan et al., 2020).

In rice, a pangenome of cultivated rice–wild rice (*O. sativa*–*O. rufipogon* Griff.) species complex through deep sequencing and de novo assembly of 66 divergent accessions was constructed (Zhao et al., 2018). In this study, intergenomic comparisons identified 23 million sequence variants in the rice genome. In maize, pangenome was characterized using 503 inbred lines and loci associated with plant developmental transitions, fitness, and adaptation traits were identified (Hirsch et al., 2014). Pangenome in wheat (*Triticum aestivum* L.) was built using 18 cultivars, which resulted in identification of 128,656 genes of which 64.3% were core genes while the remainder are variable and display presence–absence variation (Montenegro et al., 2017). In legumes, pangenome was established in soybean using seven phylogenetically and geographically representative accessions of wild soybean (*G. soja* Siebold & Zucc.), the wild relative of cultivated soybean. Analysis of the soybean pangenome identified 80% core genes. Furthermore, intergenomic comparisons identified genes associated with biotic resistance, seed composition, flowering and maturity time, and others (Li et al., 2014). Recently, another study reported pangenome of 26 representative wild and cultivated soybean selected from 2,898 globally collected soybean germplasm in terms of phylogenetic relationships and geographic distributions (Y. Liu et al., 2020). The pangenome identified large SVs and gene fusion events in soybean. The SVs identified in the study were linked to gene expression and important agronomic traits such as seed luster, seed coat pigmentation, and iron deficiency chlorosis. In the case of pigeonpea, the first pangenome based on 89 accessions was reported (Zhao et al., 2020). The study identified genes associated with important agronomic traits such as seed weight, self-fertilization, and response to disease in pigeonpea. Pangenome using eight high-quality rapeseed (*Brassica napus* L.) genomes revealed architecture and ecotype differentiation (Song et al., 2020).

Pangenome provides in-depth dissection of dispensable as well as species-specific genes identified mostly from cultivated gene pool. In order to achieve complete genetic repertoire of a given crop, diverse gene stock coming from wild species is imperative. Here comes in the concept of a super-pangenome, which represents a complete genomic variation repertoire by combining different pangenomes from all the species within a given genus (Khan et al., 2020). Deployment of the super-pangenome concept by integrating the wild side of a species with diverse genetic stock will help in tapping genetic diversity from wild species for accelerating crop improvement.

These studies provide useful insights into the genetic variability, population structure, and diversity of important crop species that could be used for crop improvement programs (Tao et al., 2019). As sequencing and resequencing costs are continuously decreasing, sequence-based allele discovery has become more prevalent. Systematic application of genome-

wide sequence information in support of crop improvement as translational genomics for agriculture will accelerate the precision of crop breeding cycle (Bohra et al., 2020; Varshney et al., 2015). The genomic resources developed in crop species will facilitate the dissection of complex traits and identification and exploitation of SNPs and SVs associated with traits of interest (Thudi et al., 2021). Furthermore, knowledge of resequencing and pangenomes and super-pangenomes would provide information on an untapped pool of diversity for easy access in breeding resulting in genetic improvement of crop species to meet future food demands.

2.1.3 | Genome sequence variations, gene prediction, and functional inferences

The major challenge for crop improvement is increasing the productivity while reducing yield losses that result from various biotic and abiotic stresses under climate change scenarios (Palit et al., 2020b). Therefore, the major aim of genomic studies in crop plants has been the identification of key regulatory genes and the active pathways associated with plant architecture and crop yields. Intensive sequence-level characterization of a chromosomal region and cloning reveals the presence of novel genes of unknown function (Jaganathan et al., 2020). Unique alleles, the makeup of alleles, variation in gene expression, signature sequences, among other, are important contributors to phenotypic diversity within and between species. Therefore, a 5G breeding approach for bringing in the much-needed disruptive changes to crop improvement has been proposed by Varshney et al. (2020). Unlimited genomics resources, such as SNPs and genome-wide SVs, are made available through sequencing and resequencing of diverse germplasm in different crops (Thudi et al., 2021). These resources would facilitate genotyping the landraces and breeding material for identification of candidate genes and diagnostic markers for the traits of interest in crop species. For example, genome sequencing of rice subspecies SN265 (*O. sativa* L. subsp. *japonica* Kato), R99 (*O. sativa* L. subsp. *indica* Kato), and resequencing of a total of 151 recombinant inbred lines generated from the cross between these parents revealed 1.7 million SNPs. Analysis of the data revealed yield and quality-associated loci and the involvement of a candidate gene, *DEP1*, in determining panicle length (X. Li et al., 2018). Similarly, the genome assembly of a maize small-kernel inbred line derived from tropical landrace provides insights into 80,614 polymorphic SVs across 521 diverse lines (N. Yang et al., 2019). Further dissection through map-based cloning of a major effect quantitative trait locus controlling kernel weight revealed the underlying candidate gene, *ZmBARELY ANY MERISTEM1d*, that provides a target for increasing maize yields.

In the case of legumes, for example, the genome of cultivated peanut (*Arachis hypogaea* L.) provided insight into legume karyotypes, polyploid evolution, and crop domestication. The QTL for seed size and testa color have been mapped to the peanut reference genome (Zhuang et al., 2019). Comprehensive analysis of resequencing data has enhanced our understanding of complex traits and provided candidate genes for several agronomic and disease-resistance traits for use in crop breeding (Varshney et al., 2019a). For instance, resequencing of 302 wild and cultivated soybean accessions revealed 230 selective sweeps and 162 selected copy number variants and correlated 96 of the 230 selected regions with oil QTL and fatty acid biosynthesis genes (Zhou et al., 2015). Similarly, in common bean (*Phaseolus vulgaris* L.), GWAS on a panel of 192 genotypes revealed candidate genes for flowering time variation (Raggi et al., 2019). In chickpea, using whole-genome resequencing data from 132 chickpea lines, GWAS identified four genetic regions containing 38 SNPs significantly associated with yield and yield-related traits (Y. Li et al., 2018). Furthermore, increase in prediction accuracy has been demonstrated in this study by incorporating results from GWAS in genomic selection. In the case of pigeonpea, markers for seed protein content were developed from whole-genome resequencing data from four pigeonpea lines demonstrating the potential of genomics (Obala et al., 2019).

With copious amount of genomics information in model and crop species, comparative functional analysis of genomics data facilitates gaining new insights toward exploring new genes and traits for potential application in crop improvement. For example, X. Yang et al. (2019) demonstrated the use of comparative genomics in evolutionary mechanisms and gene function in crassulacean acid metabolism plants. Similarly, based on genome information of pigeonpea, Kawashima et al. (2016) reported cloning of a *Phakopsora pachyrhizi* resistance gene *CcRpp1* (*Cajanus cajan* resistance against *Phakopsora pachyrhizi* 1) from pigeonpea and showed that *CcRpp1* confers resistance to *P. pachyrhizi* (causing soybean rust) in soybean. In summary, genomics resources have an enormous scope in modernizing crop breeding programs to deliver next generation varieties.

2.2 | Transcriptomics

Transcriptomics explains the conceptual changes encompassing not only in the genome itself but also the process by which the information contained in the genome is used by the cell and to discover the flow of biological information from the genome to the cell. To begin with, efforts have been focused on the development of complementary DNA libraries, generation of expressed sequence tags, gene expression analysis, and the *in silico* mining of functional infor-

mation from expressed sequence tags data sets even before genome sequences were available (Varshney et al., 2009). Initial gene expression studies relied on low-throughput methods. However, a RNA sequencing approach provides higher coverage and greater resolution of transcriptome dynamics when compared with Sanger sequencing and microarray-based methods (Garg et al., 2019).

Transcriptome sequencing is being widely used in studying plant responses to various stresses as well as its growth and development. In addition, transcriptome sequencing has been applied for various functional genomics purposes such as gene expression profiling, genome annotation, and the discovery of noncoding RNA (Morozova & Marra, 2008). Broad availability of NGS technologies led to a paradigm shift in molecular breeding of crop species and are expected to directly detect epigenetic modifications on native DNA and to allow whole-transcript sequencing without the need for genome assembly (van Dijk et al., 2018). Several transcriptome assemblies have been generated for major crops such as rice (Tian et al., 2015), wheat (Jia et al., 2018), and maize (Zhang et al., 2019a) to aid in elucidating molecular regulation of candidate genes for different traits at different stages of growth and development of the plant under stress and control conditions. In legumes, for example, hybrid comprehensive assemblies were generated in the case of pigeonpea (Kudapa et al., 2012) and chickpea (Kudapa et al., 2014) by analyzing sequencing data from three different platforms (Sanger, FLX/454, and Illumina). In a different study, transcriptome analysis under elevated CO₂ concentrations identified stress responsive candidate genes and pathways mainly involved in sugar and starch metabolism, chlorophyll, and secondary metabolites biosynthesis (Palit et al., 2020a). Gene expression profiling data from developed transcriptome assemblies enabled identification of candidate genes associated with different traits of interest and stress response.

In addition to identifying candidate genes for the traits of interest and stress response, understanding how underlying genome information translates into specific phenotypes at key developmental stages is crucial. For this, information on gene expression patterns across different plant developmental stages and organs covering entire plant life cycle is required. Gene expression atlases (GEAs) allow a thorough survey of the entire transcriptional landscape, revealing genome-wide gene activity in different tissues of several model and crop plants. In rice, GEA was developed from 39 tissues collected throughout the life cycle of the rice plant from two cultivars, Zhenshan 97 and Minghui 63 (L. Wang et al., 2010). This study provided a versatile resource for associating transcriptomics to the developmental process and understanding the regulatory process by tracing the expression profiles of individual genes. Similarly, in maize, 18 representative maize tissues capturing important aspects of maize development were targeted for GEA resulting in identification and

characterization of genes and pathways underlying plant growth and development (Sekhon et al., 2013). In legumes, for example, the RNA Seq-Atlas in soybean provided a record of high-resolution gene expression in a set of 14 diverse tissues and identification of candidate genes involved in seed development, nodule formation, and seed filling (Severin et al., 2010). Similarly, in chickpea, 27 tissues from five major developmental stages were used to construct a comprehensive *Cicer arietinum* Gene Expression Atlas (Kudapa et al., 2018), which identified significant differences in gene expression patterns contributing to the process of flowering, nodulation, seed and root development. In pigeonpea, 30 tissues representing developmental stages from germination to senescence were used to generate a *Cajanus cajan* Gene Expression Atlas (Pazhamala et al., 2017), which provided candidate genes involved in specific developmental processes and to understand the well-orchestrated growth and developmental process in pigeonpea. Similarly, *Arachis hypogaea* (groundnut) Gene Expression Atlas has been very useful to investigate complex regulatory networks, namely gravitropism and photomorphogenesis, seed development, allergens, and oil biosynthesis in groundnut (Sinha et al., 2020). These transcriptomic resources should be able to provide insights into the molecular mechanisms of growth and development, high yields, stress responses among many other traits, ultimately assisting in development of improved crop varieties.

2.3 | Proteomics

Transcriptomes ultimately come to fruition through translation to build proteins that, in combination, form the complex proteomes of different tissue types and different stages of plant development. The correlation of abundance of transcripts and proteins have been analyzed in many plant systems and notably during germination, seed development, and responses to stress. This necessitates the direct study of proteins to fully understand gene expression. The use of mass spectrometry for protein identification and protein relative or absolute quantitation, referred to as proteomics, enables the study of large sets of cellular proteins that constitutes these proteomes. Proteomics approaches have transitioned from being descriptive to become highly useful for data validation and integration with other omics approaches, providing information on biological processes and stress tolerance mechanisms that can be applied in crop breeding programs (J. Hu et al., 2015). Data-dependent ‘shotgun’ proteome sampling strategies enable large datasets of relatively quantified differences between crop varieties to be assessed but are typically limited to a small number of comparisons. Meanwhile, selected reaction-monitoring strategies allows targeted quantitation of known proteins of interest over much larger sample sets allowing whole recombinant inbred line, near-isogenic

line, and double-haploid populations to be screened for proteins of interest and their abundances (Jacoby et al., 2013)

The plant proteome undergoes significant changes because of the activation of stress-responsive pathways when subjected to biotic and abiotic stress. The proteins that are known to have significant involvement in abiotic stress response include heat-shock proteins, late embryogenesis abundant proteins, kinases and phosphatases, redox enzymes, secondary metabolism enzymes, osmolyte biosynthetic enzymes, photosynthesis, and carbon metabolism-related and enzymatic reactive oxygen species scavengers (see Hossain & Komatsu, 2012). Posttranslational modifications of proteins are also essential features of plant response to environment and are critical for plant phenotypes associated with stress tolerances (Millar et al., 2019). Protein abundance itself is only a proxy for function, rate of protein synthesis and degradation, age of proteins, and age-associated features of protein function is also critical to defining proteomes (Nelson & Millar, 2015). Several proteome maps have been developed in the model plants and crops, for instance, *Arabidopsis* (Baerenfeller et al., 2008), rice (Helmy et al., 2011), wheat (Duncan et al., 2017), barrel clover (*Medicago truncatula* Gaertn.), *Lotus japonicus* (Regel) K. Larsen, and soybean (see Ramalingam et al., 2015). Recently, proteomic analysis in three rice cultivars identified over 4,900 proteins and 1,309 differentially expressed proteins (Zhang et al., 2019b). This study identified eight genes encoding various metabolic proteins involved in brown planthopper (BPH) resistance in rice. Further, the study reported activation of the two-component response regulator protein (ORR22) that is crucial in early signal transduction in the resistance response against BPH through sustained promotion of salicylic acid. Also, key enzymes-lipoxygenases, dirigent proteins, and Ent-cassa-12,15-diene synthase (OsDTC1) in inheritable resistance against BPH were identified for use in breeding BPH-resistant rice cultivars (Zhang et al., 2019b). In maize, proteomic analysis under CO₂-enriched conditions resulted in identification of changes in protein abundance that were correlated to yield and related traits (Maurya et al., 2020). Reduced malondialdehyde content and antioxidant and antioxidative enzymes levels were observed in response to high CO₂. Further, more abundance of proteins related to Calvin cycle, protein synthesis assembly and degradation, defense, and redox homeostasis contributed to better growth and yield in elevated CO₂ was reported (Maurya et al., 2020). In legumes, for example, soybean, proteomics together with physiological and biochemical analysis led to identification of cross-tolerance mechanisms in response to heat and water stresses (Katam et al., 2020). The study reported elevated activities in antioxidant enzymes, such as increased ascorbate peroxidase enzyme, which restored oxidation levels and sustained soybean plants during stress. In addition, proteins such as MED37C, a probable mediator of RNA polymerase transcription II, were

elevated in response to combined heat stress and water stress levels (Katam et al., 2020).

2.4 | Metabolomics

After the establishment of transcriptomic and proteomic profiles, the next functional genomics challenge is the function of enzymes that build the complex milieu of primary and secondary photosynthetic catabolites as well as the sophisticated array of biosynthetic products that make up the metabolome. Metabolomics is the study of these small molecules in plants and the dynamic changes in their abundance on diurnal, developmental, and stress-responsive timescales (Fiehn et al., 2000; Weckwerth, 2003). Metabolomics encompasses a rapidly changing suite of technologies including mass spectrometry, multiple types of spectroscopy, and nuclear magnetic resonance spectroscopy. Metabolite pool sizes of major metabolites are of value in assessing the metabolome (Nunes-Nesi et al., 2019) but it is increasingly recognized that it is the flux through these major metabolite pools that contribute significantly to plant growth and development and stress biology (Moreira et al., 2019). Identifying rarer compounds including byproducts of signal transduction molecules, stress metabolism, and molecules that are part of plant acclimation process (Larrainzar et al., 2009) will be critical in breeding for adaptive metabolic traits. Metabolic compounds identified can also be further studied by correlation with transcriptome and proteome expression patterns. The cascading effects of gene expression on different levels of biological organization lead to the phenotype. Metabolites can directly influence the cellular physiology, thus closest to the phenotype (Guijas et al., 2018) but not directly related to the genome (Redestig & Costa, 2011). The profiling of metabolites corresponding to those of the transcripts under a specified condition or in a particular genotype allows an understanding of developmental processes and plant response to external stimuli or metabolism. This approach has been demonstrated in crops such as rice (C. Hu et al., 2014), soybean (Komatsu et al., 2011), and common bean (Hernández et al., 2007), in which up to 100 known metabolites have been shown to change in abundance based on geographic origin of the seeds and in response to flooding or nutrient limitation. Metabolic fingerprinting is used to identify metabolic signatures associated with stress responses without quantifying or identifying metabolites, for example, nuclear-magnetic-resonance-based approach for metabolic fingerprinting of 21 grass and legume cultivars (Bertram et al., 2010). Further, considering plant metabolome as the readout of their physiological status, metabolite-based GWAS has been utilized to dissect the genetic and biochemical bases of metabolism in crop plants (Luo, 2015). Metabolite-based GWAS has established a strong genotype–metabolite associations in maize and rice

(Chen et al., 2014; Dong et al., 2014; Matsuda et al., 2012; Wen et al., 2014). Metabolomics study would provide important insights that can serve as a basis for future crop improvement via metabolic engineering (see Kusano & Saito, 2012).

2.5 | Epigenomics

Epigenomics is the study of all epigenetic modifications in a cell. Epigenetic changes are heritable changes in gene expression and cellular functions as a result of DNA methylation, histone modifications, and biogenesis of noncoding RNAs without altering the underlying DNA sequence. Several studies in recent years helped to better understand the role of the epigenome in plant biotic and abiotic stress responses (see Agarwal et al., 2020). Further, comprehensive epigenomic studies of plant populations to correlate genotype–epigenotype–phenotype, and also the study of methyl QTL or epiGWAS will widen the understanding of mechanisms as well as functions of regulatory pathways in plant genomes (Yadav et al., 2018). Following the identification of key epigenetic regulators, epigenomics toward systems biology is needed to understand the dynamic and complex functional relationships at the plant systems level. Engineering epigenomes and epigenome-based predictive models will further accelerate molecular breeding programs for crop improvement.

2.6 | Single-cell omics

Most of the studies undertaken to understand plant biology were at the level of tissue, organ, or complete plant, which unraveled the biological activities and the genes involved. However, these studies could obscure the specific biological function of the individual cells or low-abundant biomolecules owing to the so-called ‘dilution effect’ (Libault et al., 2017). The unique functions of single cells could not be distinguished while making a bulk measurement of tissue. Studying cell phenotypes and behavior becomes imperative to understand developmental dynamics and response to environment in plants (Shulze et al., 2019). Over the last few years, there has been a tremendous technological advancement in terms of new imaging, miniaturization, automation, and microfluidics, thus enabling high-throughput sequencing of encapsulated single cells (Prakadan et al., 2017). Single-cell omics thus aim at identifying, quantifying, and characterizing different components of cells including transcriptome (Shulze et al., 2019), proteome (Dai & Chen, 2012; Levy & Slavov, 2018), and metabolome (de Souza et al., 2020) with spatiotemporal resolution. These high-resolution datasets provide insights on reconstruction of gene-regulatory and signaling networks driving cellular identity and function (Efremova

& Teichmann, 2020; Libault et al., 2017). Furthermore, the data generated from a single cell are incredibly useful for biological modeling and predictions (Stegle et al., 2015). However, in plants, owing to the challenges such as cell heterogeneity, existence of multiple cell states and cell walls, single-cell studies are limited unlike in animal and microbial cells (Libault et al., 2017). Nevertheless, single cell omics research is slowly gaining momentum with the emerging technological and computational advancement.

2.7 | Phenomics and high-throughput phenotyping

Phenomics is the key to exploit the gains of genomics resources. In recent years, crop phenomics has greatly evolved to generate multidimensional phenotypic data at multiple levels from cell level, organ level, plant level to population level (Dhondt et al., 2013; Houle et al., 2010; Lobos et al., 2017). It involves high-throughput, accurate, and automated measurements of phenotypic information such as plant growth, architecture, and composition at different scales. With the recent technological advances, large-scale phenotyping data acquisition and processing became possible, which remained a major bottleneck for functional genomics studies and crop breeding (Yang et al., 2020). Phenotypes, such as high-throughput shoot phenotyping, root phenotyping, canopy, leaf traits, among others, could be measured using high-throughput phenotyping platforms (Jin et al., 2020). Further sensor technologies now enable detailed recording of the environmental history of plants and, in turn, the dynamic response of crops to the environment. For example, drones or unmanned aerial vehicles, and pocketPlant3D equipped with multiple sensors, such as hyperspectral imaging as well as computed tomography imaging to targeted metabolic sensors, are used to measure traits such as leaf area index estimation, detect weeds and pathogens, and predict yield (Jin et al., 2020; Yang et al., 2020).

Major progress has been made in high-throughput phenotyping under controlled environments (Pratap et al., 2019). Application of such technology to field conditions are rapidly developing, including vision-guided robotics (Pieruschka & Schurr, 2019). High-throughput shoot and root phenotypic data were collected in several model plants and crop species under controlled conditions (Yang et al., 2020). The performance of a plant or crop is affected by multiple genes that interact with multiple environments throughout their growth and development. Advanced sensor, machine vision, and automation technology could now be used to record the crop dynamic response that could further be integrated with the sequence information (Jin et al., 2020). Since phenomics uses several types of sensors simultaneously, data acquisition in a systematic manner is also crucial beginning from exper-

imental set up to data generation and interpretation (Pratap et al., 2019). As the genomes of several model and nonmodel crop species have been sequenced, it is highly required to describe the whole-crop phenotype. This is important to link gene and QTL to crop phenotypes for dissecting key adaptive traits (Yang et al., 2020).

2.8 | QTL-omics

The greatest challenge to the agricultural research community is to be able to correlate and translate gene function to crop improvement in the field under the relevant set of environmental conditions. Predicting complex phenotypic traits from gene networks is complicated by genetic control and environmental effects among different growth and developmental processes of plants (Hammer et al., 2016). To resolve this, multidimensional analysis could provide useful information in understanding genotype–phenotype association, unlike single-data-type study designs. Integration of omics approaches leads to QTL mapping and identification of underlying genes. The QTL-omics will be an integral part, dealing with generation and analysis of large-scale multiomics data and is broadly defined as characterization of QTL using omics data (Kumawat et al., 2016). However, the biggest challenge is to integrate data from genome sequencing, transcriptomics, proteomics, metabolomics, and phenomics and to make sense of it. Otherwise, QTL-omics is one of the best approaches to capture the genetic variation present among the whole gene pool for specific quantitative traits in target environment. For example, modeling of QTL has led to the prediction of multi-genic traits such as leaf growth and nitrogen accumulation in maize using the systems biology approach (Reymond et al., 2003). Similar efforts were also made in tomato (*Solanum lycopersicum* L.) by integrating expression QTL with metabolite QTL (Schauer et al., 2006). In soybean, QTL-omics has been applied to characterize mapped QTL. In addition, this approach has also been used in novel QTL mapping and characterization (Kumawat et al., 2016).

All these omics approaches discussed above provide high-dimensional datasets on different modalities that are only discrete components for a comprehensive view of the plant system. In this regard, systems biology aims to integrate our understanding of how different biological components function to provide insight into the plant system and to develop predictive models on their response when perturbed.

3 | SYSTEMS BIOLOGY

One of the major goals of crop biology research is to maximize yield and reduce losses resulting from various stress factors. As the problem to be tackled is multifaceted, so is the

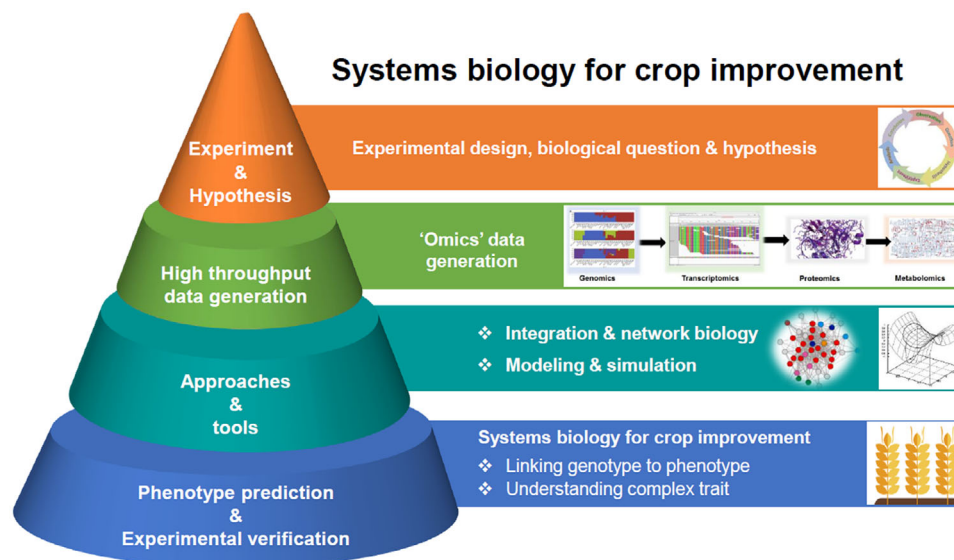


FIGURE 2 A schematical representation of the systems biology approach for crop improvement. Systems biology provides great potential for sustainable agriculture by understanding complexity of multiple traits bridging the genotype–phenotype gap. It is a strategic approach where it precisely studies the response of each level of biological organization and aims at understanding the complexity of the system as a whole through data integration coupled with mathematical modeling to gain predictive abilities over key agricultural traits

solution multidisciplinary. Understanding the cellular response at each level of the organization has been possible because of the advances in technology that have led to the generation of a huge amount of data from genome, transcriptome, proteome, metabolome, and high-throughput phenome that is routine now. However, most of the times, these data have been studied independently until recently. Integration of transcriptomics, proteomics, and metabolomics would greatly facilitate identification and dissection of complex plant regulatory networks (Urano et al., 2010). In this regard, systems biology emerged as a promising multidisciplinary research field that integrates large omics datasets coupled with well-designed mathematical models to confirm hypothesis and predict biological systems (Figure 2; Hong et al., 2019; Sauer et al., 2007). This provides a more holistic understanding of system-wide response during growth, development, and stress adaptation, critical for next-generation breeding of climate-ready crops. However, the first step to a systems biology approach is to devise hypotheses based on prior knowledge. This becomes the basis for a systems biology experimental design and is the most critical step (Pinu et al., 2019). An overview of the various steps involved in a systems biology approach and its application in crop research have been provided in the following sections.

3.1 | Data integration

The major challenge in integrating omics data remains with the processing, scaling, and analyzing the multidimensional

dataset to extract meaningful biological inferences. Integration and analysis of the datasets generated from multiple platforms involve data acquisition, preprocessing, appropriate normalization, and integration into a single matrix. This integrated dataset is generally subjected to multivariate analysis and looked for strong correlations among the biological entities. Genes, proteins, and metabolites with similar patterns were then classified into clusters (Redestig & Costa, 2011; Smilde et al., 2009). Most of the statistical methodologies include dimensionality reduction and studying the coexpressed clusters among the different data measured on the same samples. This is based on the assumption that biological entities showing similar expression patterns across the same samples have hypothetical functional relationships (González et al., 2012). Several platforms to integrate multi-dimensional omics datasets are available such as mixOmics, OnPLS modeling, Integromics, sparse Multi-Block Partial Least Squares, and COVAIN (see Misra et al., 2019; Sun & Weckwerth, 2012). In this regard, PANOMICS platform provides integration of complex omics datasets generated from genomics, epigenomics, transcriptomics, proteomics, post-translational modifications proteomics, metabolomics, and phenomics (Weckwerth et al., 2020). Recently, a systematic multiomics data integration approach, different methodologies, software tools, web applications, and databases for plant systems biology have been proposed (Jamil et al., 2020; Pinu et al., 2019).

These integrative multiomics studies have been used to identify disease mechanisms for improved prognostic and predictive marker identification that reflect molecular pathways

TABLE 1 Some examples of crop research studies using an integrated omics approach

Study no.	Plant species	Study	Omics approaches	Reference
1	Rice	Plant response to ozone	Transcriptomics, proteomics, metabolomics	Cho et al., 2008
2		Physiological and nutritional quality		Galland et al., 2017
3	Maize	Bt and glyphosate resistant		Barros et al., 2010
4		Carotenoid biosynthesis		Decoutelle et al., 2015
5		Nitrogen use efficiency		Amiour et al., 2012
6	Soybean	Primary metabolism regulation in response to <i>Rhizoctonia foliar blight</i> disease	Transcriptomics, metabolomics	Copley et al., 2017
7	Common bean	Characterizing variability	Transcriptomics, proteomics, metabolomics	Mensack et al., 2010
8	Tomato	Primary and secondary metabolism		Balcke et al., 2017
9	Sesame	Drought stress	Transcriptomics, metabolomics	You et al., 2019
10	Berry	Drought stress	Transcriptomics, proteomics, metabolomics	Ghan et al., 2015
11	Pepper	Fruit development	Transcriptomics, proteomics	Liu et al., 2019

in humans (Eddy et al., 2020; Hasin et al., 2017). Similarly, in crop plants, these approaches enable the study of plant metabolism and understand the molecular mechanisms underlying plant phenotypes with potential agronomic importance. Light-specific metabolic and regulatory signatures were identified using transcriptomics, metabolomics, and genome-scale *in silico* modeling in rice (Lakshmanan et al., 2015). Recently, transcriptomics, proteomics, and metabolomics data were analyzed to complement information that provided insights into the fertility transition mechanisms in a pigeonpea thermosensitive male sterile line for its potential use in two-line hybrid breeding (Pazhamala et al., 2020). In another study, how flavonoid and isoflavonoid metabolism alters in response to ethylene and abscisic acid treatment was studied in soybean leaves by integrating proteomics and metabolomics (Gupta et al., 2018). Table 1 provides a few of the recent studies conducted in crop plants by integrating multiomics data.

Furthermore, multiomics data provides a link between phenotype and genome variation to offer new data layers for genomics-based predictions (Azodi et al., 2020; Weckwerth et al., 2020). In crops such as maize, multiple omics data were integrated into prediction models for improved prediction accuracy (Azodi et al., 2020; Guo et al., 2016; Schrag et al., 2018; Xu et al., 2017). Multiomics data are being increasingly used for phenotypic prediction as it is not restricted to the

genome but a result of biological regulation in response to the environment (Acharjee et al., 2016; Li et al., 2019).

3.2 | Network biology

Cells respond to various genetic and environmental changes through biological processes that are regulated at multiple levels, both transcriptional and translational. Expression of genes are regulated through gene-to-gene interaction, epigenetic modification, mutations, transcription factors, and other mechanisms. Most of the plant response and adaptation to stress are specifically controlled by regulatory networks (Gehan et al., 2015; Urano et al., 2010). Reconstruction of pathways and networks using transcriptome, proteome, and metabolome data can help understand these regulatory networks and their functional interaction among the biological entities (Moreno-Risueno et al., 2010). Interaction among biological entities to carry out cellular functions can be represented as networks and graphs, elucidating biological relationships among genes, proteins, and metabolites (Weckwerth, 2011; Weckwerth et al., 2004). Briefly, omics data are appropriately normalized to obtain a similarity matrix and, subsequently, an adjacency matrix generated is transformed into an undirected graph or a network abstraction.

(Langfelder & Horvath, 2008; Langfelder et al., 2013; Redestig & Costa, 2011). Principles, methods, and tools of network inference for exploring biological details, evolutionary origin, and understanding the network structure for predicting biological functions can be found in recent reports (Hao et al., 2016; Lee et al., 2015; Saint-Antoine & Singh, 2020). Network-based approaches are limited by current knowledge as well as predicted relationships between biological variables, for instance, Bayesian networks. Several platforms to integrate multidimensional omics datasets are available (Misra et al., 2019). Further, different methodologies, software tools, web applications, and databases for integrating multidimensional omics datasets have been reported (Pinu et al., 2019).

A system can thus be mathematically represented as a set of nodes linked by edges, where nodes are the biological entity and edges indicate their interaction or relationship, and the highly connected nodes are referred to as hub nodes (Langfelder & Horvath, 2008). Hub nodes provide stronger interactions than the nodes in the periphery of the network module. Biologically, these nodes serve as regulators and could have downstream effects on the pathways (McCormack et al., 2016). Identifying hub nodes have significant implications in detecting key components controlling or affecting simple or complex traits. Thus, networks can provide new biological insights and predict key components and their regulatory influence (Albert, 2007). For this purpose, an experimentally tested genome-scale rice gene network, RiceNet, was constructed that could accurately be used to predict gene functions in monocotyledonous species (Lee et al., 2011). Further, a barley (*Hordeum vulgare* L.) coexpressed gene network generated using transcriptome data identified gene clusters associated with response to drought stress and biogenesis of cellulose (Mochida et al., 2011). In soybean, a flowering gene network could identify the regulatory roles of *GmCOL1a* and *GmCOL1b* in flowering (Wu et al., 2019). Weighted-gene coexpression network analysis (Zhang & Horvath, 2005) was used to generate gene coexpression network to identify regulatory networks and key genes controlling seed set and size (Du et al., 2017) and nodulation and nitrogen fixation (Wu et al., 2019) in soybean, whereas pollen fertility and seed set in pigeonpea (Pazhamala et al., 2017) and acquisition of desiccation tolerance in *Boea hygrometrica* (Lin et al., 2019), among many others. On the other hand, the protein–protein interaction network was found useful in investigating complex biological activities and understanding the ways in which external signals are perceived and transduced to trigger specific plant responses (Hao et al., 2016; Struk et al., 2019). In case of Arabidopsis, an *Arabidopsis thaliana* Protein Interactome Database (Cui et al., 2007; Ding & Kihara, 2019) and a dense protein–protein interaction network of plant tricarboxylic acid cycle was generated (Zhang et al., 2017). A protein interaction network associated with salt tolerance in rice was also reported (J. Wang et al., 2013). Furthermore, metabolite–

metabolite association networks constructed based on correlation algorithms can comprehensively describe the response of the biological system to environmental perturbation (Jahagirdar & Saccenti, 2020; Jahagirdar et al., 2019; Kose et al., 2001; Rosato et al., 2018; Weckwerth et al., 2004). A study clearly demonstrated the effect of different levels of plant growth regulators and agroecosystem environment on the tomato metabolome using a metabolic network (Fatima et al., 2016). The above-mentioned network inference studies were within a single omics type; however, several valuable insights were provided through integration across different omics types. Complex network interactions in nitrogen metabolism and signaling in crop plants has been reported using integrated omics approaches (Fukushima & Kusano, 2014). RiceNet was quantitatively integrated with proteomics dataset to predict proteins involved in abiotic stress resistance namely, XA21-mediated immunity (Lee et al., 2011). In tobacco (*Nicotiana tabacum* L), coexpression gene modules and metabolite modules were integrated to identify gene–metabolite relationship. The study identified key and novel genes and potential regulators of important regulatory networks including carotenoid metabolism pathway (P. Liu et al., 2020). Similarly, Mounet et al. (2009) integrated transcriptome and metabolome data to identify subsets of genes involved in fruit development and metabolism in tomato.

In brief, global gene coexpression networks has been found to be a promising approach for studying and high-throughput prediction of specialized metabolite pathways. Thus, network biology can transform our understanding of the genetic basis of how plants respond and cope with the changing environment (Wisecaver et al., 2017). To be able to do so, it is important to have an appropriate temporal design for the acquisition of omics data that needs to be analyzed with a network biology approach. Network biology is potentially a powerful tool for modeling the cellular response to adverse environmental perturbations (McCormack et al., 2016).

3.3 | Systems modeling

The next step in network biology is the dynamic modeling that allows a comprehensive view of the gene expression shaping protein behaviors in a way to elicit metabolites in response to external triggers in plants (McCormack et al., 2016). Systems modeling primarily tests biological hypotheses and further extended to predict a system-wide response. To make a biological sense of the data, computational modeling, including dynamic simulation models and machine learning approaches, is used. In simple words, the biological system is simulated based on the hypothesis devised by the biologist, which is tested over the course of time. The simulated results are then compared with the experimental data to evaluate the hypothesis and further refined to match the

experiments. This is critical to achieving a perfect prediction of the system's behavior under dynamic conditions and time (Macklin, 2019; Muthuramalingam et al., 2019).

Four major approaches have been suggested for cellular modeling: constraint-based modeling (Price et al., 2004), metabolic modeling (Steuer et al., 2006), analysis of metabolic control (Reder, 1988), and cybernetic modeling (Kompala et al., 1984). Constraint-based modeling requires information regarding metabolic reaction and flux capacity and can also perform with low experimental data. This method has clear advantage over all other methods and has been used to reconstruct metabolic networks at the whole-cell or genome scale, which is an emerging application together with omics data to increase the predictive power (Sroka et al., 2011). This approach has been found to provide deep insights into the cellular metabolism in response to abiotic stress in rice (Muthuramalingam et al., 2019; Raman & Chandra, 2009). Contrastingly, metabolic modeling corresponds to a specific component of metabolism using equations that speculate the dynamic changes in the concentration of metabolites. In plants, metabolic modeling remains limited, as it requires kinetic variables to model networks at large scale, which is not available for most of the biochemical reactions (Muthuramalingam et al., 2019). Analysis of metabolic control, on the other hand, is one of the most widely used tools for the study of control in plants, which quantifies the response of system variables (e.g. fluxes) to small changes in system parameters (e.g. the amount or activity of the individual enzymes). The cybernetic modeling framework was originally used for macroscopic input–output models describing substrate uptake, growth, and product formation. Later, this approach has been extended for application to more complex, intracellular metabolism. The software applications and mark-up languages available to facilitate systems modeling include Cell Illustrator (Nagasaki et al., 2010), COBRAToolbox (Schellenberger et al., 2011), Acorn (Sroka et al., 2011), CellNetanalyzer (Klamt et al., 2007), among many others (see Muthuramalingam et al., 2019; Pinu et al., 2019).

Machine learning uses applied statistical and computational techniques to teach machines to extract patterns in the data including features and labels for predictive modeling (Camacho et al., 2018). Generally, there are two types of machine-learning methods: supervised and unsupervised learning (McMurray & Hollich, 2009). Supervised learning includes regression algorithm and classification algorithms techniques (random forest, multivariate regression), which are used to predict the outcome with labeled data. Unsupervised learning, on the other hand, includes clustering algorithms and association rule learning, employed for clustering data, detecting outliers, and dimensionality reduction (Mishra et al., 2019). The analytics consist of data preprocessing, modeling, and active learning to deal with the complexity and intricacies of omics data. Machine learning has emerged as a pow-

erful toolbox for integration of high-dimensional biological data, extracting systems-level insights and predicting a range of outcome (Gazestani & Lewis, 2019; van Dijk et al., 2020). In plant research, phenotype prediction and understanding genotype-to-phenotype relationship remains the fundamental goal. Predicting plant phenotypes, for example, stress response and quality traits from the multiomics data is highly required for crop improvement. In this regard, a study in maize classified DNA sequence regions into active vs. (inactive) pseudogenes, using features such as DNA methylation (Sartor et al., 2019). Another study reported identification of key genes, proteins, and metabolites that are predictive of potato (*Solanum tuberosum* L.) tuber quality from transcriptomics, proteomics, and metabolomics data. This study used random forest regression for integrating the omics data (Acharjee et al., 2016). In plant research, prediction using ensemble of neural networks is quite popular. The prediction performance of phenotype depends on the predictive modeling and analytics and is an exhaustive subject that has been reviewed recently (Kim & Tagkopoulos, 2018; Tong et al., 2020).

4 | CROP IMPROVEMENT APPLICATIONS

Systems biology provides great potential for sustainable agriculture by understanding the complexity of multiple traits bridging the genotype–phenotype gap. It can be used to model and analyze multigenic traits linked with agricultural productivity such as plant architecture, nitrogen use efficiency, water use efficiency, and abiotic and biotic stress tolerance. Plant genetics and molecular biology have been extensively contributing to the plant breeding programs. However, because of the upsurge in recent high-throughput experimental analyses and computational power, there is a transformative possibility to integrate multiple disciplines to explain any given complex trait. The availability of whole-genome sequence information, generation of omics datasets with rapidly advancing technologies, analytical tools, and software (Table 2) have made it possible to study and address abiotic stress-responsive cellular, biochemical and molecular mechanisms as well as signaling processes. For instance, well-designed mathematical models based on time series datasets allow the identification of key candidate genes for potential use in the breeding programs and thus devise a systems biology-based breeding strategy (Lavarenne et al., 2018). Furthermore, they convincingly provided a roadmap to use systems biology for future breeding programs. They inferred that the ultimate practical use of systems biology is to unravel the complex interactions governing the multigenic trait for crop improvement. Development of comprehensive models by integrating multiomics data with high throughput and precise phenotyping will ensure efficient breeding programs to improve agronomically important and

TABLE 2 A list of key tools and databases for systems biology research

Tools	Description	Link	References
Biochemical, Genetic and Genomic (BiGG) knowledge base	Reconstruction of biochemical networks; Genome-scale metabolic models (GEMs)	http://bigg.ucsd.edu	Feist et al., 2009; King et al., 2016
Systems biology markup language (SBML)	A medium for representation and exchange of biochemical network models	http://www.sbml.org/	Hucka et al., 2003
JSBML 1.0	Provides a smorgasbord of options to encode systems biology models	http://sbml.org/Software/JSBML	Rodriguez et al., 2015
VirtualPlant	A software platform to support systems biology research	http://virtualplant.bio.nyu.edu	Katari et al., 2010
Babelomics	A complete suite of web tools for functional analysis of genome-scale experiments	http://www.babelomics.org.	Al-Shahrour et al., 2006
Pathway tools version 13.0	Integrated software for pathway and genome informatics and systems biology	http://BioCyc.org/download.shtml	Karp et al., 2010
TEPIC 2	A framework for fast, accurate and versatile prediction, and analysis of TF binding from epigenetics data	https://github.com/SchulzLab/TEPIC	Schmidt et al., 2019
Mergeomics	Multidimensional data integration to identify pathogenic perturbations to biological systems	Available in Bioconductor	Shu et al., 2016
Crops in silico	An integrative platform for plant systems biology research	http://cropsinsilico.org/	Zhu et al., 2016
AtPID (<i>Arabidopsis thaliana</i> protein interactome database)	An integrative platform for plant systems biology	http://atpid.biosino.org/	Cui et al., 2007
MGS	Interaction-based simulations for integrative spatial systems biology	http://mgs.spatial-computing.org/	–
VESPER 1.5	Spatial prediction software for precision agriculture	–	Whelan et al., 2002
MetExplore	Visualization of metabolites in the context of the whole network/reactions	http://metexplore.toulouse.inra.fr/joomla3/index.php	Cottret et al., 2010
ProMeTra	Visualization and combining datasets from transcriptomics, proteomics, and metabolomics	https://omictools.com/prometra-tool	Neuweger et al., 2009
KaPPA-View	Integrates transcriptomics and metabolomics data to map pathways	http://kpv.kazusa.or.jp/kappa-view/	Tokimatsu et al., 2005

complex traits in the future. This would be critical to developing future-ready crops that can sustain and increase productivity even in marginal environments faced with biotic and abiotic stresses (Gehan et al., 2015). We anticipate the use of systems biology for devising and utilizing novel breeding approaches for crop improvement in the future.

5 | CHALLENGES AND PROSPECTS

The use of multiple omics techniques (i.e., genomics, transcriptomics, proteomics, metabolomics, epigenomics, and

single-cell omics) is becoming increasingly popular in plant sciences as sequencing cost drops and expertise rises. As a result, sequencing and resequencing information that is generated for several crops enabled the identification of novel alleles from diverse sources irrespective of the availability of the genome sequence. Furthermore, rapid advances in omics technologies provide an opportunity to generate new and informative datasets in different plant species. Integration of these genomes and functional omics data with genetic and phenotypic information in an efficient manner would lead to the identification of genes and pathways responsible for important agronomic traits. On the other hand, the huge amount of

data generated on the functional components of the cell are still underutilized even in a model crop like rice (Muthuramalingam et al., 2019). Moreover, the growth and productivity of crops in the field conditions are hindered not by individual stress but a combinatorial effect of multiple stresses, both biotic and abiotic. Hence, there is an urgent need to direct future research toward the discovery of key players, molecular networks, and models that could unravel the complex cellular and molecular functions to enhance agronomic traits in crops. In this regard, systems biology offers huge potential in crop research to revolutionize our understanding on how plants respond to growth and environmental constraints (Muthuramalingam et al., 2019).

Systems biology and integrating omics approach provides a more inclusive molecular perspective of plant biology than individual approaches. However, the integrated approach of multiple omics platforms remains an ongoing challenge because of their inherent data differences. The collection of accurate multiomics data on different molecular and functional components is most critical to systems modeling and prediction, for which there are no standard protocols shared among the global community (Kim & Tagkopoulos, 2018; Macklin, 2019). In addition, model-based integration is often restricted to well-defined model organisms (Pinu et al., 2019). Constant evolution of databases and data analysis tools would help meaningful biological interpretation of multiomics data. Various technical challenges during experimentation, data generation, integrations, handling, sharing, and interpretation are comprehensively described in Misra et al. (2019) and Macklin (2019). Hence, it is high time to form a highly connected community for plant systems biology research, instead of isolated efforts, and to develop and share resources, databases, and software tools. This is even more crucial because of the urgency to double crop yields by 2050 under uncertain climate scenarios.


ACKNOWLEDGMENTS


This review describes several studies that have been undertaken at ICRISAT as part of the Systems Biology Research Initiative from ICRISAT. RKV is thankful to Science & Engineering Research Board (SERB) of Department of Science & Technology (DST), Government of India for providing the J C Bose National Fellowship (SB/S9/Z-13/2019) and Bill & Melinda Gates Foundation and CGIAR Research Program on Grain Legumes and Dryland Cereals (GLDC) for supporting the research activities at CEGSB. HK acknowledges the SERB of DST, Government of India for providing the SERB Women Excellence Award (SB/WEA-01/2017). ICRISAT is a member of CGIAR Consortium.


CONFLICT OF INTEREST

The authors declare no conflicts of interest.


ORCID

Lekha T. Pazhamala  <https://orcid.org/0000-0002-0271-7481>

Himabindu Kudapa  <https://orcid.org/0000-0003-0778-9959>

Wolfram Weckwerth  <https://orcid.org/0000-0002-9719-6358>

A. Harvey Millar  <https://orcid.org/0000-0001-9679-1473>

Rajeev K. Varshney  <https://orcid.org/0000-0002-4562-9131>

REFERENCES

- Acharjee, A., Kloosterman, B., Visser, R. G., & Maliepaard, C. (2016). Integration of multi-omics data for prediction of phenotypic traits using random forest. *BMC Bioinformatics*, 17, 363–373. <https://doi.org/10.1186/s12859-016-1043-4>
- Agarwal, G., Kudapa, H., Ramalingam, A., Choudhary, D., Sinha, P., Garg, V., Singh, V. K., Patil, G. B., Pandey, M. K., Nguyen, H. T., Guo, B., Sunkar, R., Niederhuth, C. E., & Varshney, R. K. (2020). Epigenetics and epigenomics: Underlying mechanisms, relevance, and implications in crop improvement. *Functional & Integrative Genomics*, 20, 739–761. <https://doi.org/10.1007/s10142-020-00756-7>
- Albert, R. (2007). Network inference, analysis, and modeling in systems biology. *The Plant Cell*, 19, 3327–3338. <https://doi.org/10.1105/tpc.107.054700>
- Al-Shahrouh, F., Minguez, P., Tárraga, J., Montaner, D., Alloza, E., Vaquerizas, J. M., Conde, L., Blaschke, C., Vera, J., & Dopazo, J. (2006). BABELOMICS: A systems biology perspective in the functional annotation of genome-scale experiments. *Nucleic Acids Research*, 34, W472–W476. <https://doi.org/10.1093/nar/gkl172>
- Amiour, N., Imbaud, S., Clément, G., Agier, N., Zivy, M., Valot, B., Balliau, T., Armengaud, P., Quilleré, I., Cañas, R., Tercet-Laforgue, T., & Tercet-Laforgue, T. (2012). The use of metabolomics integrated with transcriptomic and proteomic studies for identifying key steps involved in the control of nitrogen metabolism in crops such as maize. *Journal of Experimental Botany*, 63, 5017–5033. <https://doi.org/10.1093/jxb/ers186>
- Azodi, C. B., Pardo, J., VanBuren, R., de Los Campos, G., & Shiu, S. H. (2020). Transcriptome-based prediction of complex traits in maize. *The Plant Cell*, 32, 139–151. <https://doi.org/10.1105/tpc.19.00332>
- Baerenfaller, K., Grossmann, J., Grobei, M. A., Hull, R., Hirsch-Hoffmann, M., Yalovsky, S., Zimmermann, P., Grossniklaus, U., Gruissem, W., & Baginsky, S. (2008). Genome-scale proteomics reveals *Arabidopsis thaliana* gene models and proteome dynamics. *Science*, 320, 938–941. <https://doi.org/10.1126/science.1157956>
- Balcke, G. U., Bennowitz, S., Bergau, N., Athmer, B., Henning, A., Majovsky, P., Jiménez-Gómez, J. M., Hoehenwarter, W., & Tissier, A. (2017). Multi-omics of tomato glandular trichomes reveals distinct features of central carbon metabolism supporting high productivity of specialized metabolites. *The Plant Cell*, 29, 960–983. <https://doi.org/10.1105/tpc.17.00060>
- Barros, E., Lezar, S., Anttonen, M. J., Van Dijk, J. P., Röhligh, R. M., Kok, E. J., & Engel, K. H. (2010). Comparison of two GM maize varieties with a near-isogenic non-GM variety using transcriptomics, proteomics and metabolomics. *Plant Biotechnology Journal*, 8, 436–451. <https://doi.org/10.1111/j.1467-7652.2009.00487.x>

- Barutcu, A. R., Fritz, A. J., Sayyed, K. Z., van Wijnen, A. J., Lian, J. B., Stein, J. L., Nickerson, J. A., Imbalzano, A. N., & Stein, G. S. (2015). C-ing the genome: A compendium of chromosome conformation capture methods to study higher-order chromatin organization. *Journal of Cellular Physiology*, *231*, 1097–14652. <https://doi.org/10.1002/jcp.25062>
- Bertram, H. C., Weisbjerg, M. R., Jensen, C. S., Pedersen, M. G., Didion, T., Petersen, B. O., Duus, J. Ø., Larsen, M. K., & Nielsen, J. H. (2010). Seasonal changes in the metabolic fingerprint of 21 grass and legume cultivars studied by nuclear magnetic resonance-based metabolomics. *Journal of Agricultural and Food Chemistry*, *58*, 4336–4341. <https://doi.org/10.1021/jf904321p>
- Bohra, A., Jha, U. C., Godwin, I. D., & Varshney, R. K. (2020). Genomic interventions for sustainable agriculture. *Plant Biotechnology Journal*, *18*, 2388–2405. <https://doi.org/10.1111/pbi.13472>
- Camacho, D. M., Collins, K. M., Powers, R. K., Costello, J. C., & Collins, J. J. (2018). Next-generation machine learning for biological networks. *Cell*, *173*, 1581–1592. <https://doi.org/10.1016/j.cell.2018.05.015>
- Chen, W., Gao, Y., Xie, W., Gong, L., Lu, K., Wang, W., Li, Y., Liu, X., Zhang, H., Dong, H., Zhang, W., Zhang, L., Yu, S., Wang, G., Lian, X., & Luo, J. (2014). Genome-wide association analyses provide genetic and biochemical insights into natural variation in rice metabolism. *Nature Genetics*, *46*, 714–721. <https://doi.org/10.1038/ng.3007>
- Cho, K., Shibato, J., Agrawal, G. K., Jung, Y. H., Kubo, A., Jwa, N. S., Tamogami, S., Satoh, K., Kikuchi, S., Higashi, T., Kimura, S., Saji, H., Tanaka, Y., Iwahashi, H., Masuo, Y., & Rakwal, R. (2008). Integrated transcriptomics, proteomics, and metabolomics analyses to survey ozone responses in the leaves of rice seedling. *Journal of Proteome Research*, *7*, 2980–2998. <https://doi.org/10.1021/pr800128q>
- Copley, T. R., Aliferis, K. A., Kliebenstein, D. J., & Jabajim, S. H. (2017). An integrated RNAseq-1 H NMR metabolomics approach to understand soybean primary metabolism regulation in response to *Rhizoctonia foliar blight* disease. *BMC Plant Biology*, *17*, 84. <https://doi.org/10.1186/s12870-017-1020-8>
- Cottret, L., Wildridge, D., Vinson, F., Barrett, M. P., Charles, H., Sagot, M. F., & Jourdan, F. (2010). MetExplore: A web server to link metabolomic experiments and genome-scale metabolic networks. *Nucleic Acids Research*, *38*, W132–W137. <https://doi.org/10.1093/nar/gkq312>
- Cui, J., Li, P., Li, G., Xu, F., Zhao, C., Li, Y., Yang, Z., Wang, G., Yu, Q., Li, Y., & Shi, T. (2007). AtPID: Arabidopsis thaliana protein interaction database—An integrative platform for plant systems biology. *Nucleic Acids Research*, *36*, D999–D1008. <https://doi.org/10.1093/nar/gkm844>
- Dai, S., & Chen, S. (2012). Single-cell-type proteomics: Toward a holistic understanding of plant function. *Molecular & Cellular Proteomics*, *11*, 1622–1630. <https://doi.org/10.1074/mcp.R112.021550>
- de Souza, L. P., Borghi, M., & Fernie, A. (2020). Plant single-cell metabolomics—Challenges and perspectives. *International Journal of Molecular Sciences*, *21*, 8987. <https://doi.org/10.3390/ijms21238987>
- Decourcelle, M., Perez-Fons, L., Baulande, S., Steiger, S., Couvelard, L., Hem, S., Zhu, C., Capell, T., Christou, P., Fraser, P., & Sandmann, G. (2015). Combined transcript, proteome, and metabolite analysis of transgenic maize seeds engineered for enhanced carotenoid synthesis reveals pleiotropic effects in core metabolism. *Journal of Experimental Botany*, *66*, 3141–3150. <https://doi.org/10.1093/jxb/erv120>
- Dhondt, S., Wuyts, N., & Inzé, D. (2013). Cell to whole-plant phenotyping: The best is yet to come. *Trends in Plant Science*, *18*, 428–439. <https://doi.org/10.1016/j.tplants.2013.04.008>
- Ding, Z., & Kihara, D. (2019). Computational identification of protein-protein interactions in model plant proteomes. *Scientific Reports*, *9*, 8740. <https://doi.org/10.1038/s41598-019-45072-8>
- Dong, X., Chen, W., Wang, W., Zhang, H., Liu, X., & Luo, J. (2014). Comprehensive profiling and natural variation of flavonoids in rice. *Journal of Integrative Plant Biology*, *56*, 876–886. <https://doi.org/10.1111/jipb.12204>
- Du, J., Wang, S., He, C., Zhou, B., Ruan, Y. L., & Shou, H. (2017). Identification of regulatory networks and hub genes controlling soybean seed set and size using RNA sequencing analysis. *Journal of Experimental Botany*, *68*, 1955–1972. <https://doi.org/10.1093/jxb/erw460>
- Duncan, O., Trösch, J., Fenske, R., Taylor, N. L., & Millar, A. H. (2017). Resource: Mapping the *Triticum aestivum* proteome. *The Plant Journal*, *89*, 601–616. <https://doi.org/10.1111/tpj.13402>
- Eddy, S., Mariani, L. H., & Kretzler, M. (2020). Integrated multi-omics approaches to improve classification of chronic kidney disease. *Nature Reviews Nephrology*, *16*, 657–668. <https://doi.org/10.1038/s41581-020-0286-5>
- Efremova, M., & Teichmann, S. A. (2020). Computational methods for single-cell omics across modalities. *Nature Methods*, *17*, 14–17. <https://doi.org/10.1038/s41592-019-0692-4>
- Fatima, T., Sobolev, A. P., Teasdale, J. R., Kramer, M., Bunce, J., Handa, A. K., & Mattoo, A. K. (2016). Fruit metabolite networks in engineered and non-engineered tomato genotypes reveal fluidity in a hormone and agroecosystem specific manner. *Metabolomics*, *12*, 103. <https://doi.org/10.1007/s11306-016-1037-2>
- Feist, A. M., Herrgård, M. J., Thiele, I., Reed, J. L., & Palsson, B. Ø. (2009). Reconstruction of biochemical networks in microorganisms. *Nature Reviews Microbiology*, *7*, 129–143. <https://doi.org/10.1038/nrmicro1949>
- Fiehn, O., Kopka, J., Dörmann, P., Altmann, T., Trethewey, R. N., & Willmitzer, L. (2000). Metabolite profiling for plant functional genomics. *Nature Biotechnology*, *18*, 1157–1161. <https://doi.org/10.1038/81137>
- Fukushima, A., & Kusano, M. (2014). A network perspective on nitrogen metabolism from model to crop plants using integrated ‘omics’ approaches. *Journal of Experimental Botany*, *65*, 5619–5630. <https://doi.org/10.1093/jxb/eru322>
- Galland, M., He, D., Lounifi, I., Arc, E., Clément, G., Balzergue, S., & Granier, F. (2017). An integrated “multi-omics” comparison of embryo and endosperm tissue-specific features and their impact on rice seed quality. *Frontiers in Plant Science*, *8*, 1984. <https://doi.org/10.3389/fpls.2017.01984>
- Garg, V., Khan, A. W., Kudapa, H., Kale, S. M., Chitkineeni, A., Qiwei, S., Sharma, M., Li, C., Zhang, B., Xin, L., Kishor, P. B. K., & Varshney, R. K. (2019). Integrated transcriptome, small RNA and degradome sequencing approaches provide insights into Ascochyta blight resistance in chickpea. *Plant Biotechnology Journal*, *17*, 914–931. <https://doi.org/10.1111/pbi.13026>
- Gazestani, V. H., & Lewis, N. E. (2019). From genotype to phenotype: Augmenting deep learning with networks and systems biology. *Current Opinion in Systems Biology*, *15*, 68–73. <https://doi.org/10.1016/j.coisb.2019.04.001>
- Gehan, M. A., Greenham, K., Mockler, T. C., & McClung, C. R. (2015). Transcriptional networks—Crops, clocks, and abiotic stress. *Current*

- Opinion in Plant Biology*, 24, 39–46. <https://doi.org/10.1016/j.pbi.2015.01.004>
- Ghan, R., Van Sluyter, S. C., Hochberg, U., Degu, A., Hopper, D. W., Tillet, R. L., & Cramer, G. R. (2015). Five omic technologies are concordant in differentiating the biochemical characteristics of the berries of five grapevine (*Vitis vinifera* L.) cultivars. *BMC Genomics*, 16, 946. <https://doi.org/10.1186/s12864-015-2115-y>
- Giani, A. M., Gallo, G. R., Gianfranceschi, L., & Formenti, G. (2020). Long walk to genomics: History and current approaches to genome sequencing and assembly. *Computational and Structural Biotechnology Journal*, 18, 9–19. <https://doi.org/10.1016/j.csbj.2019.11.002>
- González, I., Lê Cao, K. A., Davis, M. J., & Déjean, S. (2012). Visualising associations between paired ‘omics’ data sets. *BioData Mining*, 5, 19. <https://doi.org/10.1186/1756-0381-5-19>
- Guijas, C., Montenegro-Burke, J. R., Warth, B., Spilker, M. E., & Siuzdak, G. (2018). Metabolomics activity screening for identifying metabolites that modulate phenotype. *Nature Biotechnology*, 36, 316–320. <https://doi.org/10.1038/nbt.4101>
- Guo, Z., Magwire, M. M., Basten, C. J., Xu, Z., & Wang, D. (2016). Evaluation of the utility of gene expression and metabolic information for genomic prediction in maize. *Theoretical and Applied Genetics*, 129, 2413–2427. <https://doi.org/10.1007/s00122-016-2780-5>
- Gupta, R., Min, C. W., Kramer, K., Agrawal, G. K., Rakwal, R., Park, K. H., Wang, Y., Finkemeier, I. & Kim, S. T. (2018). A multi-omics analysis of *Glycine max* leaves reveals alteration in flavonoid and isoflavonoid metabolism upon ethylene and abscisic acid treatment. *Proteomics*, 18, 1700366. <https://doi.org/10.1002/pmic.201700366>
- Hammer, G., Messina, C., van Oosterom, E., Chapman, S., Singh, V., Borrell, A., Jordan, D. R., & Cooper, M. (2016). Molecular breeding for complex adaptive traits: How integrating crop ecophysiology and modeling can enhance efficiency. In X. Yin & P. C. Struik (Eds.), *Crop systems biology* (pp. 147–162). Springer.
- Hao, T., Peng, W., Wang, Q., Wang, B., & Sun, J. (2016). Reconstruction and application of protein–protein interaction network. *International Journal of Molecular Sciences*, 17, 907. <https://doi.org/10.3390/ijms17060907>
- Hasin, Y., Seldin, M., & Lusic, A. (2017). Multi-omics approaches to disease. *Genome Biology*, 18, 83. <https://doi.org/10.1186/s13059-017-1215-1>
- Helmy, M., Tomita, M., & Ishihama, Y. (2011). OryzaPG-DB: Rice proteome database based on shotgun proteogenomics. *BMC Plant Biology*, 11, 63. <https://doi.org/10.1186/1471-2229-11-63>
- Hernández, G., Ramírez, M., Valdés-López, O., Tesfaye, M., Graham, M. A., Czechowski, T., Schlereth, A., Wandrey, M., Erban, A., Cheung, F., Wu, H. C., Lara, M., Town, C. D., Kopka, J., Udvardi, M. K., & Vance, C. P. (2007). Phosphorus stress in common bean: Root transcript and metabolic responses. *Plant Physiology*, 144, 752–767. <https://doi.org/10.1104/pp.107.096958>
- Hirsch, C. N., Foerster, J. M., Johnson, J. M., Sekhon, R. S., Muttoni, G., Vaillancourt, B., Peñagaricano, F., Lindquist, E., Pedraza, M. A., Barry, K., de Leon, N., Kaepler, S. M., & Buell, C. R. (2014). Insights into the maize pan-genome and pan-transcriptome. *The Plant Cell*, 26, 121–135. <https://doi.org/10.1105/tpc.113.119982>
- Hong, W. J., Kim, Y. J., Chandran, A. K. N., & Jung, K. H. (2019). Infrastructures of systems biology that facilitate functional genomic study in rice. *Rice*, 12, 15. <https://doi.org/10.1186/s12284-019-0276-z>
- Hossain, Z., & Komatsu, S. (2012). Contribution of proteomic studies towards understanding plant heavy metal stress response. *Frontiers in Plant Science*, 3, 310. <https://doi.org/10.3389/fpls.2012.00310>
- Houle, D., Govindaraju, D. R., & Omholt, S. (2010). Phenomics: The next challenge. *Nature Reviews Genetics*, 11, 855–866. <https://doi.org/10.1038/nrg2897>
- Hu, C., Shi, J., Quan, S., Cui, B., Kleessen, S., Nikoloski, Z., Tohge, T., Alexander, D., Guo, L., Lin, H., Wang, J., Cui, X., Rao, J., Luo, Q., Zhao, X., Fernie, A. R., & Zhang, D. (2014). Metabolic variation between *japonica* and *indica* rice cultivars as revealed by non-targeted metabolomics. *Scientific Reports*, 4, 5067. <https://doi.org/10.1038/srep05067>
- Hu, J., Rampitsch, C., & Bykova, N. V. (2015). Advances in plant proteomics toward improvement of crop productivity and stress resistance. *Frontiers in Plant Science*, 6, 209. <https://doi.org/10.3389/fpls.2015.00209>
- Hucka, M., Finney, A., Sauro, H., Bolouri, H., Doyle, J. C., Kitano, H., Arkin, A. P., Bornstein, B. J., Bray, D., Cornish-Bowden, A., Cuelar, A. A., Dronov, S., Gilles, E. D., Ginkel, M., Gor, V., Goryanin, I. I., Hedley, W. J., Hodgman, T. C., Hofmeyr, J. H., ... SBML Forum. (2003). The systems biology markup language (SBML): A medium for representation and exchange of biochemical network models. *Bioinformatics*, 19, 524–531. <https://doi.org/10.1093/bioinformatics/btg015>
- Jacoby, R. P., Millar, A. H., & Taylor, N. L. (2013). Application of selected reaction monitoring mass spectrometry to field-grown crop plants to allow dissection of the molecular mechanisms of abiotic stress tolerance. *Frontiers in Plant Science*, 4, 20. <https://doi.org/10.3389/fpls.2013.00020>
- Jaganathan, D., Bohra, A., Thudi, M., & Varshney, R. K. (2020). Fine mapping and gene cloning in the post-NGS era: Advances and prospects. *Theoretical and Applied Genetics*, 133, 1791–1810. <https://doi.org/10.1007/s00122-020-03560-w>
- Jahagirdar, S., & Saccenti, E. (2020). On the use of correlation and MI as a measure of metabolite–metabolite association for network differential connectivity analysis. *Metabolites*, 10, 171. <https://doi.org/10.3390/metabo10040171>
- Jahagirdar, S., Suarez-Diez, M., & Saccenti, E. (2019). Simulation and reconstruction of metabolite–metabolite association networks using a metabolic dynamic model and correlation based algorithms. *Journal of Proteome Research*, 18, 1099–1113. <https://doi.org/10.1021/acs.jproteome.8b00781>
- Jamil, I. N., Remali, J., Azizan, K. A., Muhammad, N. N. A., Arita, M., Goh, H. H., & Aizat, W. M. (2020). Systematic multi-omics integration (MOI) approach in plant systems biology. *Frontiers in Plant Science*, 11, 944. <https://doi.org/10.3389/fpls.2020.00944>
- Jia, M., Guan, J., Zhai, Z., Geng, S., Zhang, X., Mao, L., & Li, A. (2018). Wheat functional genomics in the era of next generation sequencing: An update. *The Crop Journal*, 6, 7–14. <https://doi.org/10.1016/j.cj.2017.09.003>
- Jiao, Y., Zhao, H., Ren, L., Song, W., Zeng, B., Guo, J., Wang, B., Liu, Z., Chen, J., Li, W., Zhang, M., Xie, S., & Lai, J. (2012). Genome-wide genetic changes during modern breeding of maize. *Nature Genetics*, 44, 812–815. <https://doi.org/10.1038/ng.2312>
- Jin, X., Zarco-Tejada, P., Schmidhalter, U., Reynolds, M. P., Hawkesford, M. J., Varshney, R. K., Yang, T., Nie, C., Li, Z., Ming, B., Xiao, Y., Xie, Y. & Li, S. (2020). High-throughput estimation of crop traits: A review of ground and aerial phenotyping platforms. *IEEE Geoscience and Remote Sensing Magazine*, 9, 200–231. <https://doi.org/10.1109/MGRS.2020.2998816>
- Karp, P. D., Paley, S. M., Krummenacker, M., Latendresse, M., Dale, J. M., & Lee, T. J., Kaipa, P., Gilham, F., Spaulding, A., Popescu, L.,

- Altman, T., Paulsen, I., Keseler, I. M., & Caspi, R. (2010). Pathway Tools version 13.0: Integrated software for pathway/genome informatics and systems biology. *Briefings in Bioinformatics*, *11*, 40–79. <https://doi.org/10.1093/bib/bbp043>
- Katam, R., Shokri, S., Murthy, N., Singh, S. K., Suravajhala, P., Khan, M. N., Bahmani, M., Sakata, K., & Reddy, K. R. (2020). Proteomics, physiological, and biochemical analysis of cross tolerance mechanisms in response to heat and water stresses in soybean. *PLoS ONE*, *5*, e0233905. <https://doi.org/10.1371/journal.pone.0233905>
- Katari, M. S., Nowicki, S. D., Aceituno, F. F., Nero, D., Kelfer, J., Thompson, L. P., & Coruzzi, G. M. (2010). VirtualPlant: A software platform to support systems biology research. *Plant Physiology*, *152*, 500–515. <https://doi.org/10.1104/pp.109.147025>
- Kawashima, C. G., Guimarães, G. A., Nogueira, S. R., MacLean, D., Cook, D. R., Steuernagel, B., Baek, J., Bouyioukos, C., Melo Bdo, V., Tristão, G., de Oliveira, J. C., Rauscher G., Mittal, S., Panichelli, L., Bacot, K., Johnson, E., Iyer, G., Tabor, G., Wulff, B. B., ... Brommonschenkel, S. H. (2016). A pigeonpea gene confers resistance to Asian soybean rust in soybean. *Nature Biotechnology*, *34*, 661. <https://doi.org/10.1038/nbt.3554>
- Kersey, P. J. (2019). Plant genome sequences: Past, present, future. *Current Opinion in Plant Biology*, *48*, 1–8. <https://doi.org/10.1016/j.pbi.2018.11.001>
- Khan, A. W., Garg, V., Roorkiwal, M., Golicz, A. A., Edwards, D., & Varshney, R. K. (2020). Super-Pangenome by integrating the wild side of a species for accelerated crop improvement. *Trends in Plant Science*, *25*, 148–158. <https://doi.org/10.1016/j.tplants.2019.10.012>
- Kim, M., & Tagkopoulos, I. (2018). Data integration and predictive modeling methods for multi-omics datasets. *Molecular Omics*, *14*, 8–25. <https://doi.org/10.1039/C7MO00051K>
- King, Z. A., Lu, J., Dräger, A., Miller, P., Federowicz, S., Lerman, J. A., Ebrahim, A., Palsson, B. O., & Lewis, N. E. (2016). BIGG Models: A platform for integrating, standardizing and sharing genome-scale models. *Nucleic Acids Research*, *44*, D515–D522. <https://doi.org/10.1093/nar/gkv1049>
- Klamt, S., Saez-Rodriguez, J., & Gilles, E. D. (2007). Structural and functional analysis of cellular networks with CellNetAnalyzer. *BMC Systems Biology*, *1*, 2. <https://doi.org/10.1186/1752-0509-1-2>
- Komatsu, S., Yamamoto, A., Nakamura, T., Nouri, M. Z., Nanjo, Y., Nishizawa, K., & Furukawa, K. (2011). Comprehensive analysis of mitochondria in roots and hypocotyls of soybean under flooding stress using proteomics and metabolomics techniques. *Journal of Proteome Research*, *10*, 3993–4004. <https://doi.org/10.1021/pr2001918>
- Kompala, D. S., Ramkrishna, D., & Tsao, G. T. (1984). Cybernetic modeling of microbial growth on multiple substrates. *Biotechnology and Bioengineering*, *26*, 1272–1281. <https://doi.org/10.1002/bit.260261103>
- Kose, F., Weckwerth, W., Linke, T., & Fiehn, O. (2001). Visualizing plant metabolomic correlation networks using clique–metabolite matrices. *Bioinformatics*, *17*, 1198–1208. <https://doi.org/10.1093/bioinformatics/17.12.1198>
- Kudapa, H., Azam, S., Sharpe, A. G., Taran, B., Li, R., Deonovic, B., Cameron, C., Farmer, A. D., Cannon, S. B., & Varshney, R. K. (2014). Comprehensive transcriptome assembly of chickpea (*Cicer arietinum* L.) using Sanger and next generation sequencing platforms: Development and applications. *PLoS ONE*, *9*, e86039. <https://doi.org/10.1371/journal.pone.0086039>
- Kudapa, H., Bharti, A. K., Cannon, S. B., Farmer, A. D., Mulaosmanovic, B., Kramer, R., Bohra, A., Weeks, N. T., Crow, J. A., Tuteja, R., Shah, T., Dutta, S., Gupta, D. K., Singh, A., Gaikwad, K., Sharma, T. R., May, G. D., Singh, N. K., & Varshney, R. K. (2012). A comprehensive transcriptome assembly of pigeonpea (*Cajanus cajan* L.) using Sanger and second-generation sequencing platforms. *Molecular Plant*, *5*, 1020–1028. <https://doi.org/10.1093/mp/ssr111>
- Kudapa, H., Garg, V., Chitikineni, A., & Varshney, R. K. (2018). The RNA-Seq-based high resolution gene expression atlas of chickpea (*Cicer arietinum* L.) reveals dynamic spatio-temporal changes associated with growth and development. *Plant, Cell & Environment*, *41*, 2209–2225. <https://doi.org/10.1111/pce.13210>
- Kumawat, G., Gupta, S., Ratnaparkhe, M. B., Maranna, S., & Satpute, G. K. (2016). QTLomics in soybean: A way forward for translational genomics and breeding. *Frontiers in Plant Science*, *7*, 1852. <https://doi.org/10.3389/fpls.2016.01852>
- Kusano, M., & Saito, K. (2012). Role of metabolomics in crop improvement. *Journal of Plant Biochemistry and Biotechnology*, *21*, 24–31. <https://doi.org/10.1007/s13562-012-0131-4>
- Lakshmanan, M., Lim, S. H., Mohanty, B., Kim, J. K., Ha, S. H., & Lee, D. Y. (2015). Unraveling the light-specific metabolic and regulatory signatures of rice through combined in silico modeling and multiomics analysis. *Plant Physiology*, *169*, 3002–3020. <https://doi.org/10.1093/plphys/kpv104>
- Lam, H. M., Xu, X., Liu, X., Chen, W., Yang, G., Wong, F. L., Li, M. W., He, W., Qin, N., Wang, B., Li, J., Jian, M., Wang, J., Shao, G., Wang, J., Sun, S. S., & Zhang, G. (2010). Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. *Nature Genetics*, *42*, 1053. <https://doi.org/10.1038/ng.715>
- Langfelder, P., & Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics*, *9*, 559. <https://doi.org/10.1186/1471-2105-9-559>
- Langfelder, P., Mischel, P. S., & Horvath, S. (2013). When is hub gene selection better than standard meta-analysis? *PLoS ONE*, *8*, e61505. <https://doi.org/10.1371/journal.pone.0061505>
- Larrainzar, E., Wienkoop, S., Scherling, C., Kempa, S., Ladrera, R., Arrese-Igor, C., Weckwerth, W., & González, E. M. (2009). Carbon metabolism and bacteroid functioning are involved in the regulation of nitrogen fixation in *Medicago truncatula* under drought and recovery. *Molecular Plant-Microbe Interactions*, *22*, 1565–1576. <https://doi.org/10.1094/MPMI-22-12-1565>
- Lavarenne, J., Guyomarç'h, S., Sallaud, C., Gantet, P., & Lucas, M. (2018). The spring of systems biology-driven breeding. *Trends in Plant Science*, *23*, 706–720. <https://doi.org/10.1016/j.tplants.2018.04.005>
- Lee, I., Seo, Y. S., Coltrane, D., Hwang, S., Oh, T., Marcotte, E. M., & Ronald, P. C. (2011). Genetic dissection of the biotic stress response using a genome-scale gene network for rice. *Proceedings of the National Academy of Sciences*, *108*, 18548–18553. <https://doi.org/10.1073/pnas.1110384108>
- Lee, T., Kim, H., & Lee, I. (2015). Network-assisted crop systems genetics: Network inference and integrative analysis. *Current Opinion in Plant Biology*, *24*, 61–70. <https://doi.org/10.1016/j.pbi.2015.02.001>
- Levy, E., & Slavov, N. (2018). Single cell protein analysis for systems biology. *Essays in Biochemistry*, *62*, 595–605. <https://doi.org/10.1042/EBC20180014>
- Li, X., Wu, L., Wang, J., Sun, J., Xia, X., Geng, X., Wang, X., Xu, Z., & Xu, Q. (2018). Genome sequencing of rice subspecies and genetic analysis of recombinant lines reveals regional yield- and quality-associated loci. *BMC Biology*, *16*, 102. <https://doi.org/10.1186/s12915-018-0572-x>

- Li, Y., Ruperao, P., Batley, J., Edwards, D., Khan, T., Colmer, T. D., Pang, J., Siddique, K. H. M., & Sutton, T. (2018). Investigating drought tolerance in chickpea using genome-wide association mapping and genomic selection based on whole-genome resequencing data. *Frontiers in Plant Science*, *19*, 190. <https://doi.org/10.3389/fpls.2018.00190>
- Li, Y. H., Zhou, G., Ma, J., Jiang, W., Jin, L. G., Zhang, Z., Guo, Y., Zhang, J., Sui, Y., Zheng, L., Zhang, S. S., Zuo, Q., Shi, X. H., Li, Y. F., Zhang, W. K., Hu, Y., Kong, G., Hong, H. L., Tan, B., ... Qiu, L. J. (2014). De novo assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits. *Nature Biotechnology*, *32*, 1045–1052. <https://doi.org/10.1038/nbt.2979>
- Li, Z., Gao, N., Martini, J. W., & Simianer, H. (2019). Integrating gene expression data into genomic prediction. *Frontiers in Genetics*, *10*, 126. <https://doi.org/10.3389/fgene.2019.00126>
- Libault, M., Pingault, L., Zogli, P., & Schiefelbein, J. (2017). Plant systems biology at the single-cell level. *Trends in Plant Science*, *22*, 949–960. <https://doi.org/10.1016/j.tplants.2017.08.006>
- Lin, C. T., Xu, T., Xing, S. L., Zhao, L., Sun, R. Z., Liu, Y., Moore, J. P. & Deng, X. (2019). Weighted gene co-expression network analysis (WGCNA) reveals the hub role of protein ubiquitination in the acquisition of desiccation tolerance in *Boea hygrometrica*. *Plant and Cell Physiology*, *60*, 2707–2719. <https://doi.org/10.1093/pcp/pcz160>
- Liu, P., Luo, J., Zheng, Q., Chen, Q., Zhai, N., Xu, S., Xu, Y., Jin, L., Xu, G., Lu, X., Xu, G., Wang, G., Shao, J., Xu H. M., Cao, P., Zhou, H., Wang, X., & Xu, G. (2020). Integrating transcriptome and metabolome reveals molecular networks involved in genetic and environmental variation in tobacco. *DNA Research*, *27*, dsaa006. <https://doi.org/10.1093/dnares/dsaa006>
- Liu, Y., Du, H., Li, P., Shen, Y., Peng, H., Liu, S., Zhou, G. A., Zhang, H., Liu, Z., Shi, M., Huang, X., Li, Y., Zhang, M., Wang, Z., Zhu, B., Han B., Liang, C., & Tian, Z. (2020). Pan-genome of wild and cultivated soybeans. *Cell*, *182*, 162–176.E13. <https://doi.org/10.1016/j.cell.2020.05.023>
- Liu, Z., Lv, J., Zhang, Z., Li, H., Yang, B., Chen, W., Dai, X., Li, X., Yang, S., Liu, L., Ou, L., Ma, Y., & Zou, X. (2019). Integrative transcriptome and proteome analysis identifies major metabolic pathways involved in pepper fruit development. *Journal of Proteome Research*, *18*, 982–994. <https://doi.org/10.1021/acs.jproteome.8b00673>
- Lobos, G. A., Camargo, A. V., del Pozo, A., Araus, J. L., Ortiz, R., & Doonan, J. H. (2017). Plant phenotyping and phenomics for plant breeding. *Frontiers in Plant Science*, *8*, 2181. <https://doi.org/10.3389/fpls.2017.02181>
- Luo, J. (2015). Metabolite-based genome-wide association studies in plants. *Current Opinion in Plant Biology*, *24*, 31–38. <https://doi.org/10.1016/j.pbi.2015.01.006>
- Mace, E. S., Tai, S., Gilding, E. K., Li, Y., Prentis, P. J., Bian, L., Campbell, B. C., Hu, W., Innes, D. J., Han, X., Cruickshank, A., Dai, C., Frère, C., Zhang, H., Hunt, C. H., Wang, X., Shatte, T., Wang, M., Su, Z., ... Wang, J. (2013). Whole-genome sequencing reveals untapped genetic potential in Africa's indigenous cereal crop sorghum. *Nature Communications*, *4*, 2320. <https://doi.org/10.1038/ncomms3320>
- Macklin, P. (2019). Key challenges facing data-driven multicellular systems biology. *GigaScience*, *8*, giz127. <https://doi.org/10.1093/gigascience/giz127>
- Matsuda, F., Okazaki, Y., Oikawa, A., Kusano, M., Nakabayashi, R., Kikuchi, J., Yonemaru, J., Ebana, K., Yano, M., & Saito, K. (2012). Dissection of genotype–phenotype associations in rice grains using metabolome quantitative trait loci analysis. *The Plant Journal*, *70*, 624–636. <https://doi.org/10.1111/j.1365-313X.2012.04903.x>
- Maurya, V. K., Gupta, S. K., Sharma, M., Majumder, B., Deeba, F., Pandey, N., & Pandey, V. (2020). Proteomic changes may lead to yield alteration in maize under carbon dioxide enriched condition. *3 Biotech*, *10*, 203.
- McCormack, M. E., Lopez, J. A., Crocker, T. H., & Mukhtar, M. S. (2016). Making the right connections: Network biology and plant immune system dynamics. *Current Plant Biology*, *5*, 2–12. <https://doi.org/10.1016/j.cpb.2015.10.002>
- McMurray, B., & Hollich, G. (2009). Core computational principles of language acquisition: Can statistical learning do the job? Introduction to Special Section. *Developmental Science*, *12*, 365–368. <https://doi.org/10.1111/j.1467-7687.2009.00821.x>
- Mensack, M. M., Fitzgerald, V. K., Ryan, E. P., Lewis, M. R., Thompson, H. J., & Brick, M. A. (2010). Evaluation of diversity among common beans (*Phaseolus vulgaris* L.) from two centers of domestication using 'omics' technologies. *BMC Genomics*, *11*, 686. <https://doi.org/10.1186/1471-2164-11-686>
- Millar, A. H., Heazlewood, J. L., Giglione, C., Holdsworth, M. J., Bachmair, A., & Schulze, W. X. (2019). The scope, functions and dynamics of post-translational protein modifications. *Annual Review of Plant Biology*, *70*, 119–151. <https://doi.org/10.1146/annurev-arplant-050718-100211>
- Mishra, B., Kumar, N., & Mukhtar, M. S. (2019). Systems biology and machine learning in plant–pathogen interactions. *Molecular Plant-Microbe Interactions*, *32*, 45–55. <https://doi.org/10.1094/MPMI-08-18-0221-FI>
- Misra, B. B., Langefeld, C., Olivier, M., & Cox, L. A. (2019). Integrated omics: Tools, advances and future approaches. *Journal of Molecular Endocrinology*, *62*, R21–R45. <https://doi.org/10.1530/JME-18-0055>
- Mochida, K., Uehara-Yamaguchi, Y., Yoshida, T., Sakurai, T., & Shinzaki, K. (2011). Global landscape of a co-expressed gene network in barley and its application to gene discovery in Triticeae crops. *Plant and Cell Physiology*, *52*, 785–803. <https://doi.org/10.1093/pcp/pcr035>
- Montenegro, J. D., Golicz, A. A., Bayer, P. E., Hurgobin, B., Lee, H., Chan, C. K. K., Visendi, P., Lai, K., Doležel, J., Batley, J., & Edwards, D. (2017). The pangenome of hexaploid bread wheat. *The Plant Journal*, *90*, 1007–1013. <https://doi.org/10.1111/tjp.13515>
- Moreira, T. B., Shaw, R., Luo, X., Ganguly, O., Kim, H. S., Coelho, L. G., Cheung, C. Y. M., & Rhys Williams, T. C. (2019). A genome-scale metabolic model of soybean (*Glycine max*) highlights metabolic fluxes in seedlings. *Plant Physiology*, *180*, 1912–1929. <https://doi.org/10.1104/pp.19.00122>
- Moreno-Risueno, M. A., Busch, W., & Benfey, P. N. (2010). Omics meet networks—Using systems approaches to infer regulatory networks in plants. *Current Opinion in Plant Biology*, *13*, 126–131. <https://doi.org/10.1016/j.pbi.2009.11.005>
- Morozova, O., & Marra, M. A. (2008). Applications of next-generation sequencing technologies in functional genomics. *Genomics*, *92*, 255–264. <https://doi.org/10.1016/j.ygeno.2008.07.001>
- Mounet, F., Moing, A., Garcia, V., Petit, J., Maucourt, M., Deborde, C., Bernillon, S., Le Gall, G., Colquhoun, I., Defernez, M., Giraudel, J. L., Rolin, D., Rothan, C., & Lemaire-Chamley, M. (2009). Gene and metabolite regulatory network analysis of early developing fruit tissues highlights new candidate genes for the control of tomato fruit composition and development. *Plant Physiology*, *149*, 1505–1528. <https://doi.org/10.1104/pp.108.133967>

- Muthuramalingam, P., Jeyasri, R., Krishnan, S. R., Pandian, S. T. K., Sathishkumar, R., & Ramesh, M. (2019). Integrating the bioinformatics and omics tools for systems analysis of abiotic stress tolerance in *Oryza sativa* (L.). In R. Sathishkumar, S. R. Kumar, J. Hema, & V. Baskar (Eds.), *Advances in plant transgenics: Methods and applications* (pp. 59–77), Springer.
- Nagasaki, M., Saito, A., Jeong, E., Li, C., Kojima, K., Ikeda, E., & Miyano, S. (2010). Cell Illustrator 4.0: A computational platform for systems biology. *In Silico Biology*, *10*, 5–26. <https://doi.org/10.3233/ISB-2010-0415>
- Nelson, C. J., & Millar, A. H. (2015). Protein turnover in plant biology. *Nature Plants*, *1*, 15017. <https://doi.org/10.1038/nplants.2015.17>
- Neuweger, H., Persicke, M., Albaum, S. P., Bekel, T., Dondrup, M., Hüser, A. T., Winnebal, J., Schneider, J., Kalinowski, J., & Goemann, A. (2009). Visualizing post genomics data-sets on customized pathway maps by ProMeTra—aeration-dependent gene expression and metabolism of *Corynebacterium glutamicum* as an example. *BMC Systems Biology*, *3*, 82. <https://doi.org/10.1186/1752-0509-3-82>
- Nunes-Nesi, A., Alseekh, S., de Oliveira Silva, F. M., Omranian, N., Lichtenstein, G., Mirnezhad, M., Romero González, R. R., Sabio y Garcia, J., Conte, M., Leiss, K. A., Klinkhamer, P. G. L., Nikoloski, Z., Carrari, F., & Fernie, A. R. (2019). Identification and characterization of metabolite quantitative trait loci in tomato leaves and comparison with those reported for fruits and seeds. *Metabolomics*, *15*, 46. <https://doi.org/10.1007/s11306-019-1503-8>
- Obala, J., Saxena, R. K., Singh, V. K., Kumar, C. S., Saxena, K. B., Tongoona, P., Sibiy, J., & Varshney, R. K. (2019). Development of sequence-based markers for seed protein content in pigeonpea. *Molecular Genetics and Genomics*, *294*, 57–68. <https://doi.org/10.1007/s00438-018-1484-8>
- Palit, P., Ghosh, R., Tolani, P., Tarafdar, A., Chitkineni, A., Bajaj, P., Sibiy, J., & Varshney, R. K. (2020a). Molecular and physiological alterations under elevated CO₂ concentrations in chickpea. *Plant and Cell Physiology*, *61*, 1449–1463. <https://doi.org/10.1093/pcp/pcaa077>
- Palit, P., Kudapa, H., Zougmore, R., Kholova, J., Whitbread, A., Sharma, M., & Varshney, R. K. (2020b). An integrated research framework combining genomics, systems biology, physiology, modelling and breeding for legume improvement in response to elevated CO₂ under climate change scenario. *Current Plant Biology*, *22*, 100149. <https://doi.org/10.1016/j.cpb.2020.100149>
- Pazhamala, L. T., Chaturvedi, P., Bajaj, P., Srikanth, S., Ghatak, A., Chitkineni, A., Bellaire, A., Hingane, A., Sameer Kumar, C. V., Saxena, K. B., Weckwerth, W., Saxena, R. K., & Varshney, R. K. (2020). Multiomics approach unravels fertility transition in a pigeonpea line for a two-line hybrid system. *The Plant Genome*, *13*, e20028. <https://doi.org/10.1002/tpg2.20028>
- Pazhamala, L. T., Purohit, S., Saxena, R. K., Garg, V., Krishnamurthy, L., Verdier, J., & Varshney, R. K. (2017). Gene expression atlas of pigeonpea and its application to gain insights into genes associated with pollen fertility implicated in seed formation. *Journal of Experimental Botany*, *68*, 2037–2054. <https://doi.org/10.1093/jxb/erx010>
- Pieruschka, R., & Schurr, U. (2019). Plant phenotyping: Past, present, and future. *Plant Phenomics*, *2019*, 7507131. <https://doi.org/10.34133/2019/7507131>
- Pinu, F. R., Beale, D. J., Paten, A. M., Kouremenos, K., Swarup, S., Schirra, H. J., & Wishart, D. (2019). Systems biology and multi-omics integration: Viewpoints from the metabolomics research community. *Metabolites*, *9*, 76. <https://doi.org/10.3390/metabo9040076>
- Prakadan, S. M., Shalek, A. K., & Weitz, D. A. (2017). Scaling by shrinking: Empowering single-cell 'omics' with microfluidic devices. *Nature Reviews Genetics*, *18*, 345–361. <https://doi.org/10.1038/nrg.2017.15>
- Pratap, A., Gupta, S., Nair, R. M., Gupta, S. K., Schafleitner, R., Basu, P. S., Singh, C. M., Prajapati, U., Gupta, A. K., Nayyar, H., Mishra, A. K., & Baek, K. H. (2019). Using plant phenomics to exploit the gains of genomics. *Agronomy*, *9*, 126. <https://doi.org/10.3390/agronomy9030126>
- Price, N. D., Reed, J. L., & Pálsson, B. Ø. (2004). Genome-scale models of microbial cells: Evaluating the consequences of constraints. *Nature Reviews Microbiology*, *2*, 886–897. <https://doi.org/10.1038/nrmicro1023>
- Raggi, L., Caproni, L., Carboni, A., & Negri, V. (2019). Genome-wide association study reveals candidate genes for flowering time variation in common bean (*Phaseolus vulgaris* L.). *Frontiers in Plant Science*, *10*, 962. <https://doi.org/10.3389/fpls.2019.00962>
- Ramalingam, A., Kudapa, H., Pazhamala, L. T., Weckwerth, W., & Varshney, R. K. (2015). Proteomics and metabolomics: Two emerging areas for legume improvement. *Frontiers in Plant Science*, *6*, 1116. <https://doi.org/10.3389/fpls.2015.01116>
- Raman, K., & Chandra, N. (2009). Flux balance analysis of biological systems: Applications and challenges. *Briefings in Bioinformatics*, *10*, 435–449. <https://doi.org/10.1093/bib/bbp011>
- Reder, C. (1988). Metabolic control theory: A structural approach. *Journal of Theoretical Biology*, *135*, 175–201. [https://doi.org/10.1016/S0022-5193\(88\)80073-0](https://doi.org/10.1016/S0022-5193(88)80073-0)
- Redestig, H., & Costa, I. G. (2011). Detection and interpretation of metabolite–transcript coresponses using combined profiling data. *Bioinformatics*, *27*, i357–i365. <https://doi.org/10.1093/bioinformatics/btr231>
- Reymond, M., Muller, B., Leonardi, A., Charcosset, A., & Tardieu, F. (2003). Combining quantitative trait loci analysis and an ecophysiological model to analyze the genetic variability of the responses of maize leaf growth to temperature and water deficit. *Plant Physiology*, *131*, 664–675. <https://doi.org/10.1104/pp.013839>
- Rodriguez, N., Thomas, A., Watanabe, L., Vazirabad, I. Y., Kofia, V., Gómez, H. F., Mittag, F., Matthes, J., Rudolph, J., Wrzodek, F., Netz, E., Diamantikos, N. E., Eichner, J., Keller, R., Wrzodek, C., Fröhlich, S., Lewis, N. E., Myers, C. J., Le Novère, N., ... Dräger, A., (2015). JSBML 1.0: Providing a smorgasbord of options to encode systems biology models. *Bioinformatics*, *31*, 3383–3386. <https://doi.org/10.1093/bioinformatics/btv341>
- Rosato, A., Tenori, L., Cascante, M., Carulla, P. R. D. A., dos Santos, V. A. M., & Saccenti, E. (2018). From correlation to causation: Analysis of metabolomics data using systems biology approaches. *Metabolomics*, *14*, 37. <https://doi.org/10.1007/s11306-018-1335-y>
- Saint-Antoine, M. M., & Singh, A. (2020). Network inference in systems biology: Recent developments, challenges, and applications. *Current Opinion in Biotechnology*, *63*, 89–98. <https://doi.org/10.1016/j.copbio.2019.12.002>
- Sartor, R. C., Noshay, J., Springer, N. M., & Briggs, S. P. (2019). Identification of the expressome by machine learning on omics data. *Proceedings of the National Academy of Sciences*, *116*, 18119–18125. <https://doi.org/10.1073/pnas.1813645116>
- Sauer, U., Heinemann, M., & Zamboni, N. (2007). Getting closer to the whole picture. *Science*, *316*, 550–551. <https://doi.org/10.1126/science.1142502>

- Schauer, N., Semel, Y., Roessner, U., Gur, A., Balbo, I., Carrari, F., Pleban, T., Perez-Melis, A., Bruedigam, C., Kopka, J., Willmitzer, L., Zamir, D., & Fernie, A. R. (2006). Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nature Biotechnology*, *24*, 447–454. <https://doi.org/10.1038/nbt1192>
- Schellenberger, J., Que, R., Fleming, R. M., Thiele, I., Orth, J. D., Feist, A. M., Zielinski, D. C., Bordbar, A., Lewis, N. E., Rahmani, S., Kang, J., Hyduke, D. R., & Palsson, B. Ø. (2011). Quantitative prediction of cellular metabolism with constraint-based models: The COBRA Toolbox v2.0. *Nature Protocols*, *6*, 1290–1307. <https://doi.org/10.1038/nprot.2011.308>
- Schmidt, F., Kern, F., Ebert, P., Baumgarten, N., & Schulz, M. H. (2019). TEPIC 2—An extended framework for transcription factor binding prediction and integrative epigenomic analysis. *Bioinformatics*, *35*, 1608–1609. <https://doi.org/10.1093/bioinformatics/bty856>
- Schrag, T. A., Westhues, M., Schipprack, W., Seifert, F., Thiemann, A., Scholten, S., & Melchinger, A. E. (2018). Beyond genomic prediction: Combining different types of omics data can improve prediction of hybrid performance in maize. *Genetics*, *208*, 1373–1385. <https://doi.org/10.1534/genetics.117.300374>
- Sekhon, R. S., Briskine, R., Hirsch, C. N., Myers, C. L., Springer, N. M., Buell, C. R., Kaeppler, S. M., & Kaeppler, S. M. (2013). Maize gene atlas developed by RNA sequencing and comparative evaluation of transcriptomes based on RNA sequencing and microarrays. *PLoS ONE*, *8*, e61005. <https://doi.org/10.1371/journal.pone.0061005>
- Severin, A. J., Woody, J. L., Bolon, Y. T., Joseph, B., Diers, B. W., Farmer, A. D., Muehlbauer, G. J., Nelson, R. T., Grant, D., Specht, J. E., Graham, M. A., Cannon, S. B., May, G. D., Vance, C. P., & Shoemaker, R. C. (2010). RNA-Seq atlas of *Glycine max*: A guide to the soybean transcriptome. *BMC Plant Biology*, *10*, 160. <https://doi.org/10.1186/1471-2229-10-160>
- Shu, L., Zhao, Y., Kurt, Z., Byars, S. G., Tukiainen, T., Kettunen, J., Orozco, L. D., Pellegrini, M., Lusia, A. J., Ripatti, S., Zhang, B., Inouye, M., Mäkinen, V. P., & Yang, X. (2016). Mergeomics: Multidimensional data integration to identify pathogenic perturbations to biological systems. *BMC Genomics*, *17*, 874. <https://doi.org/10.1186/s12864-016-3198-9>
- Shulse, C. N., Cole, B. J., Ciobanu, D., Lin, J., Yoshinaga, Y., Gouran, M., Turco, G. M., Zhu, Y., O'Malley, R. C., Brady, S. M., & Dickel, D. E. (2019). High-throughput single-cell transcriptome profiling of plant cell types. *Cell Reports*, *27*, 2241–2247.e4. <https://doi.org/10.1016/j.celrep.2019.04.054>
- Sinha, P., Bajaj, P., Pazhamala, L. T., Nayak, S. N., Pandey, M. K., Chitkineni, A., Huai, D., Khan, A. W., Desai, A., Jiang, H., Zhuang, W., Guo, B., Liao, B., & Varshney, R. K. (2020). *Arachis hypogaea* gene expression atlas for *fastigiata* subspecies of cultivated groundnut to accelerate functional and translational genomics applications. *Plant Biotechnology Journal*, *18*, 2187–2200. <https://doi.org/10.1111/pbi.13374>
- Smilde, A. K., Kiers, H. A., Bijlsma, S., Rubingh, C. M., & Van Erk, M. J. (2009). Matrix correlations for high-dimensional data: The modified RV-coefficient. *Bioinformatics*, *25*, 401–405. <https://doi.org/10.1093/bioinformatics/btn634>
- Song, J. M., Guan, Z., Hu, J., Guo, C., Yang, Z., Wang, S., Liu, D., Wang, B., Lu, S., Zhou, R., Xie, W. Z., Cheng, Y., Zhang, Y., Liu, K., Yang, Q. Y., Chen, L. L., & Guo, L. (2020). Eight high-quality genomes reveal pan-genome architecture and ecotype differentiation of *Brassica napus*. *Nature Plants*, *6*, 34–45. <https://doi.org/10.1038/s41477-019-0577-7>
- Sroka, J., Bieniasz-Krzywiec, Ł., Gwóźdz, S., Leniowski, D., Łacki, J., Markowski, M., Avignone-Rossa, C., Bushell, M. E., McFadden, J., & Kierzek, A. M. (2011). Acorn: A grid computing system for constraint based modeling and visualization of the genome scale metabolic reaction networks via a web interface. *BMC Bioinformatics*, *12*, 196. <https://doi.org/10.1186/1471-2105-12-196>
- Stegle, O., Teichmann, S. A., & Marioni, J. C. (2015). Computational and analytical challenges in single-cell transcriptomics. *Nature Reviews Genetics*, *16*, 133–145. <https://doi.org/10.1038/nrg3833>
- Steuer, R., Gross, T., Selbig, J., & Blasius, B. (2006). Structural kinetic modeling of metabolic networks. *Proceedings of the National Academy of Sciences*, *103*, 11868–11873. <https://doi.org/10.1073/pnas.0600013103>
- Struk, S., Jacobs, A., Sánchez Martín-Fontecha, E., Gevaert, K., Cubas, P., & Goormachtig, S. (2019). Exploring the protein–protein interaction landscape in plants. *Plant, Cell & Environment*, *42*, 387–409. <https://doi.org/10.1111/pce.13433>
- Sun, X., & Weckwerth, W. (2012). COVAIN: A toolbox for uni- and multivariate statistics, time-series and correlation network analysis and inverse estimation of the differential Jacobian from metabolomics covariance data. *Metabolomics*, *8*, 81–93. <https://doi.org/10.1007/s11306-012-0399-3>
- Tao, Y., Zhao, X., Mace, E., Henry, R., & Jordan, D. (2019). Exploring and exploiting pan-genomics for crop improvement. *Molecular Plant*, *12*, 156–169. <https://doi.org/10.1016/j.molp.2018.12.016>
- Thudi, M., Palakurthi, R., Schnable, J. C., Chitkineni, A., Dreisigacker, S., Mace, E., Srivastava, R. K., Satyavathi, C. T., Odeny, D., Tiwari, V. K., Lam, H. M., Hong, Y. B., Singh, V. K., Li, G., Xu, Y., Chen, X., Kaila, S., Nguyen, H., Sivasankar, S., ... Varshney, R. K. (2021). Genomic resources in plant breeding for sustainable agriculture. *Journal of Plant Physiology*, *257*, 153351. <https://doi.org/10.1016/j.jplph.2020.153351>
- Tian, X. J., Long, Y., Wang, J., Zhang, J. W., Wang, Y. Y., Li, W. M., Peng, Y. F., Yuan, Q. H., & Pei, X. W. (2015). De novo transcriptome assembly of common wild rice (*Oryza rufipogon* Griff.) and discovery of drought-response genes in root tissue based on transcriptomic data. *PLoS ONE*, *10*, e0131455. <https://doi.org/10.1371/journal.pone.0131455>
- Tokimatsu, T., Sakurai, N., Suzuki, H., Ohta, H., Nishitani, K., Koyama, T., Umezawa, T., Misawa, N., Saito, K., & Shibata, D. (2005). KaPPA-View. A web-based analysis tool for integration of transcript and metabolite data on plant metabolic pathway maps. *Plant Physiology*, *138*, 1289–1300. <https://doi.org/10.1104/pp.105.060525>
- Tong, H., Madison, I., Long, T. A., & Williams, C. M. (2020). Computational solutions for modeling and controlling plant response to abiotic stresses: A review with focus on iron deficiency. *Current Opinion in Plant Biology*, *57*, 8–15. <https://doi.org/10.1016/j.pbi.2020.05.006>
- Urano, K., Kurihara, Y., Seki, M., & Shinozaki, K. (2010). ‘Omics’ analyses of regulatory networks in plant abiotic stress responses. *Current Opinion in Plant Biology*, *13*, 132–138. <https://doi.org/10.1016/j.pbi.2009.12.006>
- Valliyodan, B., Brown, A. V., Wang, J., Patil, G., Liu, Y., Otyama, P. I., Nelson, R. T., Vuong, T., Song, Q., Muskat, T. A., Wagner, R., Marri, P., Reddy, S., Sessions, A., Wu, X., Grant, D., Bayer, P. E., Roorkiwal, M., Varshney, R. K., ... Nguyen, H. T. (2021). Genetic variation among 481 diverse soybean accessions, inferred

- from genomic re-sequencing. *Scientific Data*, 8, 50. <https://doi.org/10.1038/s41597-021-00834-w>
- van Dijk, A. D. J., Kootstra, G., Kruijer, W., & de Ridder, D. (2020). Machine learning in plant science and plant breeding. *iScience*, 101890. <https://doi.org/10.1016/j.isci.2020.101890>
- van Dijk, E. L., Jaszczyszyn, Y., Naquin, D., & Thermes, C. (2018). The third revolution in sequencing technology. *Trends in Genetics*, 34, 666–681. <https://doi.org/10.1016/j.tig.2018.05.008>
- Varshney, R. K., Graner, A., & Sorrells, M. E. (2005). Genomics-assisted breeding for crop improvement. *Trends in Plant Science*, 10, 621–630. <https://doi.org/10.1016/j.tplants.2005.10.004>
- Varshney, R. K., Kudapa, H., Pazhamala, L., Chitkineni, A., Thudi, M., Bohra, A., Gaur, P. M., Janila, P., Fikre, A., Kimurto, P., & Ellis, N. (2015). Translational genomics in agriculture: Some examples in grain legumes. *Critical Reviews in Plant Sciences*, 34, 169–194. <https://doi.org/10.1080/07352689.2014.897909>
- Varshney, R. K., Nayak, S. N., May, G. D., & Jackson, S. A. (2009). Next-generation sequencing technologies and their implications for crop genetics and breeding. *Trends in Biotechnology*, 27, 522–530. <https://doi.org/10.1016/j.tibtech.2009.05.006>
- Varshney, R. K., Pandey, M. K., Bohra, A., Singh, V. K., Thudi, M., & Saxena, R. K. (2019a). Toward the sequence-based breeding in legumes in the post-genome sequencing era. *Theoretical and Applied Genetics*, 132, 797–816. <https://doi.org/10.1007/s00122-018-3252-x>
- Varshney, R. K., Saxena, R. K., Upadhyaya, H. D., Khan, A. W., Yu, Y., Kim, C., Rathore, A., Kim, D., Kim, J., An, S., Kumar, V., Anuradha, G., Yamini, K. N., Zhang, W., Muniswamy, S., Kim, J. S., Penmetsa, R. V., von Wettberg E., & Datta, S. K. (2017a). Whole-genome resequencing of 292 pigeonpea accessions identifies genomic regions associated with domestication and agronomic traits. *Nature Genetics*, 49, 1082–1088. <https://doi.org/10.1038/ng.3872>
- Varshney, R. K., Shi, C., Thudi, M., Mariac, C., Wallace, J., Qi, P., Zhang, H., Zhao, Y., Wang, X., Rathore, A., Srivastava, R. K., Chitkineni, A., Fan, G., Bajaj, P., Punnuri, S., Gupta, S. K., Wang, H., Jiang, Y., Couderc, M., ... Xu, X. (2017b). Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments. *Nature Biotechnology*, 35, 969–976. <https://doi.org/10.1038/nbt.3943>
- Varshney, R. K., Sinha, P., Singh, V. K., Kumar, A., Zhang, Q., & Bennetzen, J. L. (2020). 5Gs for crop genetic improvement. *Current Opinion in Plant Biology*, 56, 190–196. <https://doi.org/10.1016/j.pbi.2019.12.004>
- Varshney, R. K., Thudi, M., Roorkiwal, M., He, W., Upadhyaya, H. D., Yang, W., Bajaj, P., Cubry, P., Rathore, A., Jian, J., Doddamani, D., Khan, A. W., Garg, V., Chitkineni, A., Xu, D., Gaur, P. M., Singh, N. P., Chaturvedi, S. K., Nadigatla, G. V. P. R., ... Liu, X. (2019b). Resequencing of 429 chickpea accessions from 45 countries provides insights into genome diversity, domestication and agronomic traits. *Nature Genetics*, 51, 857–864. <https://doi.org/10.1038/s41588-019-0401-3>
- Wang, J., Chen, L., Wang, Y., Zhang, J., Liang, Y., & Xu, D. (2013). A computational systems biology study for understanding salt tolerance mechanism in rice. *PLoS ONE*, 8, e64929. <https://doi.org/10.1371/journal.pone.0064929>
- Wang, L., Xie, W., Chen, Y., Tang, W., Yang, J., Ye, R., Liu, L., Lin, Y., Xu, C., Xiao, J., & Zhang, Q. (2010). A dynamic gene expression atlas covering the entire life cycle of rice. *The Plant Journal*, 61, 752–766. <https://doi.org/10.1111/j.1365-313X.2009.04100.x>
- Wang, W., Mauleon, R., Hu, Z., Chebotarov, D., Tai, S., Wu, Z., Li, M., Zheng, T., Fuentes, R. R., Zhang, F., Mansueto, L., Copetti, D., Sanciangco, M., Palis, K. C., Xu, J., Sun, C., Fu, B., Zhang, H., Gao, Y., Zhao, X., & Leung, H. (2018). Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature*, 557, 43–49. <https://doi.org/10.1038/s41586-018-0063-9>
- Weckwerth, W. (2003). Metabolomics in systems biology. *Annual Review of Plant Biology*, 54, 669–689. <https://doi.org/10.1146/annurev.arplant.54.031902.135014>
- Weckwerth, W. (2011). Green systems biology—From single genomes, proteomes and metabolomes to ecosystems research and biotechnology. *Journal of Proteomics*, 75, 284–305. <https://doi.org/10.1016/j.jprot.2011.07.010>
- Weckwerth, W., Ghatak, A., Bellaire, A., Chaturvedi, P., & Varshney, R. K. (2020). PANOMICS meets germplasm. *Plant Biotechnology Journal*, 18, 1507–1525. <https://doi.org/10.1111/pbi.13372>
- Weckwerth, W., Loureiro, M. E., Wenzel, K., & Fiehn, O. (2004). Differential metabolic networks unravel the effects of silent plant phenotypes. *Proceedings of the National Academy of Sciences*, 101, 7809–7814. <https://doi.org/10.1073/pnas.0303415101>
- Wen, W., Li, D., Li, X., Gao, Y., Li, W., Li, H., Liu, J., Liu, H., Chen, W., Luo, J., Yan, J., & Yan, J. (2014). Metabolome-based genome-wide association study of maize kernel leads to novel biochemical insights. *Nature Communications*, 5, 3438. <https://doi.org/10.1038/ncomms4438>
- Whelan, B. M., McBratney, A. B., & Minasny, B. (2002). Vesper 1.5—Spatial prediction software for precision agriculture. In P. C. Robert, R. H. Rust, & W. E. Larson (Eds.), *Precision Agriculture, Proceedings of the 6th International Conference on Precision Agriculture* (Vol. 179, pp. 854–867). ASA, CSSA, SSSA.
- Wisecaver, J. H., Borowsky, A. T., Tzin, V., Jander, G., Kliebenstein, D. J., & Rokas, A. (2017). A global coexpression network approach for connecting genes to specialized metabolic pathways in plants. *The Plant Cell*, 29, 944–959. <https://doi.org/10.1105/tpc.17.00009>
- Wu, F., Kang, X. K., Wang, M., Haider, W., Price, W. B., Hajek, B., & Hanzawa, Y. (2019). Transcriptome-enabled network inference revealed the GmCOL1 feed-forward loop and its roles in photoperiodic flowering of soybean. *Frontiers in Plant Science*, 10, 1221. <https://doi.org/10.3389/fpls.2019.01221>
- Xu, Y., Xu, C., & Xu, S. (2017). Prediction and association mapping of agronomic traits in maize using multiple omic data. *Heredity*, 119, 174–184. <https://doi.org/10.1038/hdy.2017.27>
- Yadav, C. B., Pandey, G., Muthamilarasan, M., & Prasad, M. (2018). Epigenetics and epigenomics of plants. In R. Varshney, M. Pandey, & A. Chitkineni (Eds.) *Plant genetics and molecular biology. Advances in biochemical engineering/biotechnology*. (Vol. 164, pp. 237–261). Springer. https://doi.org/10.1007/10_2017_51
- Yang, N., Liu, J., Gao, Q., Gui, S., Chen, L., Yang, L., Huang, J., Deng, T., Luo, J., He, L., Wang, Y., Xu, P., Peng, Y., Shi, Z., Lan, L., Ma, Z., Yang, X., Zhang, Q., Bai, M., ... Yan, J. (2019). Genome assembly of a tropical maize inbred line provides insights into structural variation and crop improvement. *Nature Genetics*, 51, 1052–1059. <https://doi.org/10.1038/s41588-019-0427-6>
- Yang, W., Feng, H., Zhang, X., Zhang, J., Doonan, J. H., Batchelor, W. D., Xiong, L., & Yan, J. (2020). Crop phenomics and high-throughput phenotyping: Past decades, current challenges, and future perspectives. *Molecular Plant*, 13, 187–214. <https://doi.org/10.1016/j.molp.2020.01.008>

- Yang, X., Liu, D., Tschaplinski, T. J., & Tuskan, G. A. (2019). Comparative genomics can provide new insights into the evolutionary mechanisms and gene function in CAM plants. *Journal of Experimental Botany*, *70*, 6539–6547. <https://doi.org/10.1093/jxb/erz408>
- You, J., Zhang, Y., Liu, A., Li, D., Wang, X., Dossa, K., Zhou, R., Yu, J., Zhang, Y., Wang, L., & Zhang, X. (2019). Transcriptomic and metabolomic profiling of drought-tolerant and susceptible sesame genotypes in response to drought stress. *BMC Plant Biology*, *19*, 267. <https://doi.org/10.1186/s12870-019-1880-1>
- Zhang, B., & Horvath, S. (2005). A general framework for weighted gene co-expression network analysis. *Statistical Applications in Genetics and Molecular Biology*, *4*, 17. <https://doi.org/10.2202/1544-6115.1128>
- Zhang, X., Wang, Y., Yan, Y., Peng, H., Long, Y., Zhang, Y., Jiang, Z., Liu, P., Zou, C., Peng, H., Pan, G., Shen, Y., & Pan, G. (2019a). Transcriptome sequencing analysis of maize embryonic callus during early redifferentiation. *BMC Genomics*, *20*, 159. <https://doi.org/10.1186/s12864-019-5506-7>
- Zhang, X., Yin, F., Xiao, S., Jiang, C., Yu, T., Chen, L., Ke, X., Zhong, Q., Cheng, Z., & Li, W. (2019b). Proteomic analysis of the rice (*Oryza officinalis*) provides clues on molecular tagging of proteins for brown planthopper resistance. *BMC Plant Biology*, *19*, 30. <https://doi.org/10.1186/s12870-018-1622-9>
- Zhang, Y., Beard, K. F., Swart, C., Bergmann, S., Krahnert, I., Nikoloski, Z., Graf, A., Ratcliffe, R. G., Sweetlove, L. J., Fernie, A. R., & Obata, T. (2017). Protein–protein interactions and metabolite channelling in the plant tricarboxylic acid cycle. *Nature Communications*, *8*, 15212. <https://doi.org/10.1038/ncomms15212>
- Zhao, J., Bayer, P. E., Ruperao, P., Saxena, R. K., Khan, A. W., Golicz, A. A., Nguyen, H. T., Batley, J., Edwards, D., & Varshney, R. K. (2020). Trait associations in the pangenome of pigeon pea (*Cajanus cajan*). *Plant Biotechnology Journal*, *18*, 1946–1954. <https://doi.org/10.1111/pbi.13354>
- Zhao, Q., Feng, Q., Lu, H., Li, Y., Wang, A., Tian, Q., Zhan, Q., Lu, Y., Zhang, L., Huang, T., Wang, Y., Fan, D., Zhao, Y., Wang, Z., Zhou, C., Chen, J., Zhu, C., Li, W., ... Huang, X. (2018). Pangenome analysis highlights the extent of genomic variation in cultivated and wild rice. *Nature Genetics*, *50*, 278–284. <https://doi.org/10.1038/s41588-018-0041-z>
- Zhou, Z., Jiang, Y., Wang, Z., Gou, Z., Lyu, J., Li, W., Yu, Y., Shu, L., Zhao, Y., Ma, Y., Fang, C., Shen, Y., Liu, T., Li C., Li, Q., Wu, M., Wang, M., Wu, Y., Dong, Y., ... Tian, Z. (2015). Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. *Nature Biotechnology*, *33*, 408–414. <https://doi.org/10.1038/nbt.3096>
- Zhu, X. G., Lynch, J. P., LeBauer, D. S., Millar, A. J., Stitt, M., & Long, S. P. (2016). Plants *in silico*: Why, why now and what? - an integrative platform for plant systems biology research. *Plant, Cell & Environment*, *39*, 1049–1057. <https://doi.org/10.1111/pce.12673>
- Zhuang, W., Chen, H., Yang, M., Wang, J., Pandey, M. K., Zhang, C., Chang, W. C., Zhang, L., Zhang, X., Tang, R., Garg, V., Wang, X., Tang, H., Chow, C. N., Wang, J., Deng, Y., Wang, D., Khan, A. W., Yang, Q., ... Varshney, R. K. (2019). The genome of cultivated peanut provides insight into legume karyotypes, polyploid evolution and crop domestication. *Nature Genetics*, *51*, 865–876. <https://doi.org/10.1038/s41588-019-0402-2>

How to cite this article: Pazhamala L. T., Kudapa H., Weckwerth W., Millar A. H., & Varshney R. K. Systems biology for crop improvement. *Plant Genome*. 2021;14:e20098. <https://doi.org/10.1002/tpg2.20098>