



## OPEN ACCESS

## EDITED BY

Lin Chen,  
Chinese Academy of Agricultural Sciences,  
China

## REVIEWED BY

Chenxia Cheng,  
Qingdao Agricultural University, China  
Junpeng Shi,  
Sun Yat-sen University, China

## \*CORRESPONDENCE

Zeyang Ma

✉ zeyangma@cau.edu.cn

RECEIVED 22 April 2024

ACCEPTED 23 May 2024

PUBLISHED 05 June 2024

## CITATION

Huo Q, Song R and Ma Z (2024) Recent advances in exploring transcriptional regulatory landscape of crops. *Front. Plant Sci.* 15:1421503. doi: 10.3389/fpls.2024.1421503

## COPYRIGHT

© 2024 Huo, Song and Ma. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Recent advances in exploring transcriptional regulatory landscape of crops

Qiang Huo, Rentao Song and Zeyang Ma\*

State Key Laboratory of Maize Bio-breeding, Frontiers Science Center for Molecular Design Breeding, Joint International Research Laboratory of Crop Molecular Breeding, National Maize Improvement Center, College of Agronomy and Biotechnology, China Agricultural University, Beijing, China

Crop breeding entails developing and selecting plant varieties with improved agronomic traits. Modern molecular techniques, such as genome editing, enable more efficient manipulation of plant phenotype by altering the expression of particular regulatory or functional genes. Hence, it is essential to thoroughly comprehend the transcriptional regulatory mechanisms that underpin these traits. In the multi-omics era, a large amount of omics data has been generated for diverse crop species, including genomics, epigenomics, transcriptomics, proteomics, and single-cell omics. The abundant data resources and the emergence of advanced computational tools offer unprecedented opportunities for obtaining a holistic view and profound understanding of the regulatory processes linked to desirable traits. This review focuses on integrated network approaches that utilize multi-omics data to investigate gene expression regulation. Various types of regulatory networks and their inference methods are discussed, focusing on recent advancements in crop plants. The integration of multi-omics data has been proven to be crucial for the construction of high-confidence regulatory networks. With the refinement of these methodologies, they will significantly enhance crop breeding efforts and contribute to global food security.

## KEYWORDS

gene regulatory network, transcription factor, multi-omics, transcriptional regulation, crop improvement

## 1 Introduction

Plant development and response to environmental stimuli rely on the precise orchestration of gene expression (Zhong et al., 2023). The rich gene expression patterns are governed by multiple regulatory mechanisms, such as gene transcription, mRNA processing, translation, and protein modifications. While gene expression is fine-tuned at different levels, transcriptional regulation is crucial and serves as the primary determinant of the cellular transcriptome (Zhong et al., 2023).

At the transcriptional level, gene expression is controlled by various factors, including transcription factors (TFs) and other DNA-binding proteins. TFs bind to specific genomic binding sites, known as cis-regulatory elements (CREs), within certain chromatin contexts. They can either activate or repress the expression of downstream target genes (Strader et al., 2022). TFs often act in a combination manner, enabling a limited number of TFs to regulate a larger set of target genes (Brkljacic and Grotewold, 2017). The coordinated action of interacting TFs (protein-protein interactions), the interactions between TFs and the promoter DNA of target genes (protein-DNA interactions), and the regulatory relationships among TFs form complex regulatory networks.

Unraveling the transcriptional regulation landscape in plants is important for improving our understanding of the regulatory principles. It allows us to understand how plants respond to internal signals and external environmental variations at the molecular level and how these changes influence plant growth and development. To implement precise genetic engineering strategies in modern breeding, manipulating key transcriptional regulators or their corresponding CREs through genetic engineering can modulate the expression of a set of functional genes or entire metabolic pathways (Grotewold, 2008). A comprehensive understanding of the regulatory networks can help to predict and mitigate potential unintended outcomes of gene editing, thereby improving the yield, nutritional quality, and resistance to diseases or environmental stresses of crop plants. For example, mutation of a target binding site in the *Ideal Plant Architecture 1 (IPA1)* promoter for an upstream TF has been reported to be able to overcome the tradeoff between the number of grains per panicle and the number of tillers in rice, leading to an increased yield (Song et al., 2022).

Significant advancements have been made in transcriptional regulation studies over the past two decades. With the advent of high-throughput DNA and protein profiling technologies, there is a growing accumulation of multi-omics data. In parallel, developing advanced computational algorithms has facilitated the integration of large-scale datasets, such as transcriptomics, epigenomics, and proteomics, enabling the reconstruction of complex regulatory networks (Depuydt et al., 2023). We are now capable of constructing more accurate network models, which contribute to a deeper understanding of gene regulation. More recently, the application of single-cell sequencing technologies has revealed the heterogeneity of transcriptional profiles at the cellular level, shedding light on the understanding of the dynamic nature of gene regulation during development and stress responses at an unprecedented resolution (Badia-i-Mompel et al., 2023).

In this review, we briefly summarize the characteristics of commonly used molecular networks. We provide an update on various transcriptional regulatory network inference approaches with multi-omics datasets, highlighting recent advances and limitations of each method. Furthermore, we outline the general downstream analyses for the reconstructed networks. Additionally, we highlight the cutting-edge progress of regulatory network studies in crop plants, with a focus on cereals, such as maize and wheat. Finally, the challenges and future directions in the field are discussed.

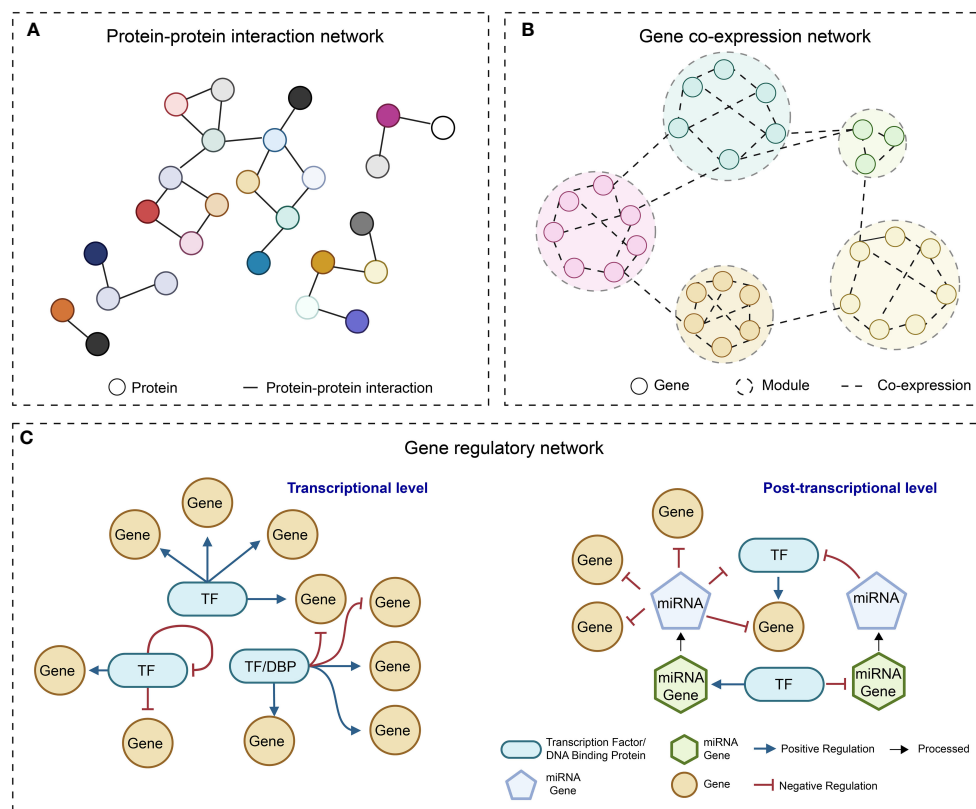
## 2 Understanding the transcriptional regulation with network-based approaches

In the post-genomic era, the accumulation of multi-omics data and the rapid progress in developing computational algorithms have empowered us to uncover the complexities of gene function and regulatory programs at a system level. The reconstruction of molecular networks, which are mainly composed of two components (nodes representing biomolecules such as proteins and nucleotides, and edges depicting the interactions between the nodes), is a straightforward approach for visualizing complex interactions and hunting for desirable genes. Most of the currently adopted molecular networks can be classified as protein-protein interaction (PPI) network (Xing et al., 2016), gene co-expression network (GCN) (Rao and Dixon, 2019), and gene regulatory network (GRN) (Van den Broeck et al., 2020) (Figure 1).

### 2.1 PPI network: depicting the interactome of proteins

Proteins that share specific biological functions are expected to be interconnected within a PPI network. Therefore, the primary purpose of PPI networks is to unravel the functions of unidentified proteins by using the annotations of known genes. In addition, the network structure information can facilitate addressing biological questions, including the identification of hub proteins, novel pathways, and evolutionary analysis of proteins of interest (von Mering et al., 2002). Noteworthy, the link between two proteins has various implications, such as altering the kinetic properties of enzymes, affecting the substrate binding affinity of effectors, and modifying the regulatory effects of TFs on their downstream target genes (Berggard et al., 2007). Given that transcription regulation depends on both TFs and their associated cofactors, PPI networks offer supplementary insights into transcription regulation (Serebreni and Stark, 2021). This additional layer of information is distinct from the link between TF and regulatory DNA elements.

Currently, only a limited number of experimental-based proteomic networks have been established in plants (Consortium, 2011). Using a yeast-two-hybrid (Y2H) mapping workflow, a comprehensive map of the phytohormone signaling network was constructed, revealing the multifaceted functions of phytohormone proteins in *Arabidopsis* (Altmann et al., 2020). Protein mass spectrometry was also used to identify protein complexes in 13 plant species (McWhite et al., 2020). Recently, Han and colleagues conducted a Y2H screening of 7,623 baits against 21,964 prey proteins in maize, resulting in the identification of 56,243 high-confidence PPIs by vigorous filtering (Han et al., 2023). Moreover, there are computational algorithms have been developed for PPI prediction (Zhang et al., 2016a). A support vector machine (SVM) model has been trained to generate a Protein-Protein Interaction Database for Maize (PPIM), covering ~ 2,700,000 interactions among ~ 14,000 proteins (Zhu et al., 2016). Yang and coworkers



**FIGURE 1**  
Simple diagrams of different molecular networks. (A) protein-protein interaction network; (B) Gene co-expression network; (C) Gene regulatory network at transcriptional level and post-transcriptional level.

have established a comprehensive database (PlaPPISite) for 13 plant interactomes by collecting experimentally validated PPIs and computational predictions (Yang et al., 2020).

Despite these significant advances, well-explored plant PPI maps remain limited. Current data shows that the available PPI datasets cover ~ 12,000 genes in *Arabidopsis*, ~ 40 genes in Soybean, and ~ 300 genes in Rice based on the Biogrid database (accessed on 20 April 2024) (Oughtred et al., 2021). Therefore, it is important to establish a larger quantity of high-quality PPI networks through coordinated efforts by the research communities in the future.

## 2.2 GCN: a useful tool to predict gene function

With the rapid accumulation of transcriptomics data, such as gene expression microarrays and RNA-seq data, GCNs are frequently employed to elucidate the connection between genes and to cluster a large number of genes that exhibit similar expression patterns (Stuart et al., 2003). GCNs represent indirect connections without considering directionality. They are typically generated by a weighted network construction approach followed by hierarchical clustering to identify smaller co-expression modules (Langfelder and Horvath, 2008). While we can use prior knowledge

of TF-coding genes to assign the directionality from TFs to their target genes, the directionality between two TFs always remains unknown in the GCN. Despite the lack of causal regulatory links, mounting evidence suggests that GCNs are efficient in predicting the specific biological functions of unknown genes by the “guilt-by-association” principle (Wolfe et al., 2005) and in identifying hub genes that exhibit high connectivity with other genes and may have important regulatory roles (Lin et al., 2019).

To pinpoint regulatory or functional genes involved in specific biological processes, functional modules associated with various pathways or traits are partitioned from large GCNs and annotated by Gene Ontology (GO) terms enrichment analysis. For example, a GCN was constructed and divided into 25 modules in wheat. These modules were annotated and connected to the spatiotemporal progression during wheat endosperm development (Pfeifer et al., 2014); Co-expression modules were also identified for secondary biosynthetic pathways in tea plants (Tai et al., 2018).

As GCNs inherently lack information regarding regulatory relationships among co-expressed genes, it is necessary to combine co-expression analysis with additional complementary data sources, such as cis-regulatory data. Integration approaches can enhance the reliability of GCNs for capturing true biological relevance from network connections. By integration of co-expression data, cis-regulatory elements, and conserved DNA motifs, Vandepoel and coworkers were able to accurately link

many unknown genes to specific biological functions, such as the E2F pathway in *Arabidopsis* (Vandepoele et al., 2009).

### 2.3 GRN: a primary approach for investigating regulatory codes

In plants and other organisms, TFs regulate structural genes and TF-coding genes with high context specificity (Badia-i-Mompel et al., 2023). GRN analysis serves as a robust tool for delineating the regulatory relationships between a single or a set of TFs with distinct functions and their downstream target genes in specific cellular and environmental conditions. It has also shown its value in identifying key regulator TFs, regulatory connections between genes and pathways, and in formulating testable functional and regulatory hypotheses (Van den Broeck et al., 2020).

GRNs can be classified into two groups based on their objectives: context-dependent GRNs and comprehensive untargeted GRNs (Depuydt et al., 2023). The majority of GRN studies have been designed to elucidate the network wiring that underlies specific developmental processes or responses to particular environmental conditions. For example, a Bayesian-based network analysis was used to identify multiple genes associated with the *SHOOT MERISTEMLESS* TF gene and to predict their roles in shoot apical meristem development (Scofield et al., 2018). Borrill and colleagues integrated time-series data in wheat and identified several hub genes, including the well-known senescence regulator *NAM-A1*, which regulates the expression of senescence-related genes within the network (Borrill et al., 2019). Zander and coworkers generated a GRN model to predict the cross-talk in the jasmonic acid (JA) signaling pathway and to discover novel components involved in the JA regulatory mechanism (Zander et al., 2020). Furthermore, known and novel candidate TFs were identified associated with water-deficit responses and xylem development plasticity using integrative network analysis in rice (Reynoso et al., 2022).

While context-dependent GRN studies often provide high-resolution information on the specific biological process under investigation, untargeted GRN approaches, despite having lower resolution, are able to capture a broader range of biological processes under various conditions. Untargeted GRNs are typically generated using extensive datasets without focusing on only one specific biological question. Instead, they have been used to establish a database resource or test novel algorithms. For example, Zhou and colleagues collected extensive transcriptome datasets to create coexpression-based GRNs in maize (Zhou et al., 2020). Recently, several resource articles have been published, such as MaizeNetome (Feng et al., 2023), Wheat-RegNet (Tang et al., 2023), and wGRN (Chen et al., 2023b). To introduce more different context-specificities, it is common to incorporate a lot of complementary datasets from various tissues, treatments, and developmental stages. Moreover, the integration of additional omics layers, such as trait-association results, can provide further evidence for the hypotheses drawn from the transcriptome and identify more accurate candidates for the following experimental

validation (Kim et al., 2023). Nevertheless, although these GRNs are very large, containing millions of edges, they are not saturated yet.

### 2.4 Inference of gene network in the single-cell era

Single-cell omics technologies, particularly single-cell RNA sequencing (scRNA-seq), provide comprehensive insights into the transcription landscape of diverse plant tissues, surpassing conventional bulk sequencing methods (Rhee et al., 2019). As gene regulation principally takes place in individual cells, inferring regulatory networks based on single-cell data is more effective than using bulk data. It predicts interactions based on expression within the same cells rather than averages (Chen et al., 2019). Moreover, the increased resolution of single-cell omics data allows to capture the cell type- or state-specific GRNs (Aibar et al., 2017).

Current single-cell assays are limited in their ability to detect all transcripts in every cell, often capturing fewer than 5,000 genes per cell. Therefore, specialized tools have been developed to handle this data sparsity (Hao et al., 2024). Common network inference methods designed for scRNA-seq data include SCODE (Matsumoto et al., 2017), PID (Chan et al., 2017), Inferelator (Jackson et al., 2020), and SCENIC/SCENIC+ (Aibar et al., 2017; Van de Sande et al., 2020; Bravo Gonzalez-Blas et al., 2023). These methods vary in their underlying models for linking regulator TFs to target genes. SCENIC first identifies regulatory relationships based on co-expressed genes using GENIE3 (Huynh-Thu et al., 2010) or GRNBoost2 (Moerman et al., 2019), and then refines the connections by considering TF binding motifs on promoter regions. The defined “regulons” consist of co-expressed genes enriched for the CREs to which the regulatory TF binds. Finally, the workflow identifies cells where these regulons are active (Aibar et al., 2017). However, the lack of validated and formatted TF-DNA binding data for most plant species hinders the application of these methods with plant scRNA-seq data.

Single-cell technologies now allow for the quantification of many other modalities, such as scATAC-seq (Buenrostro et al., 2013). GRN methods have been developed to combine the data from multiple modalities (Jiang et al., 2022a; Alanis-Lobato et al., 2024), or alternatively, networks can be constructed separately with each modality and then integrated together (Jansen et al., 2019). Nevertheless, unlike bulk sequencing technologies, which capture a higher number of transcripts, the sparsity inherent in single-cell data may result in biased estimations of gene expression correlations (van Dijk et al., 2018). We expect these challenges to be addressed through enhanced sequencing depths and more sophisticated bioinformatics methodologies to effectively manage data with limited counts (Sekula et al., 2020; Song et al., 2023).

Currently, compared to PPI network and GCN, GRN has emerged as a favored tool for predicting essential regulators and gene expression alterations in response to environmental stimuli and intrinsic signals (Gupta et al., 2022). In some articles, broadly defined GRNs can be formed by the connections between regulatory elements that regulate the transcriptional and translational

processes. Such elements, including TFs, splicing factors, and microRNAs, could be incorporated into the modeling of GRNs (Lai et al., 2016; Carthew, 2021) (Figure 1). In the following sections, we will use the term “GRN” to refer to the network that abstracts the directed relationships between TFs and their target genes in the context of transcriptional regulation and emphasize studies related to GRNs.

experimental techniques and computational inference methods (van der Sande et al., 2023). Here, we adopt a classification based on the source data types of the regulatory link, dividing the networks into three categories: physical, functional, and integrative regulatory networks (Figure 2) (Marbach et al., 2012b). The selection of methods for constructing regulatory networks depends on the specific research goals and the availability of relevant data.

### 3 Reconstruction of transcriptional regulatory networks with multi-omics data

GRNs describe the relationship between target genes and their upstream regulator TFs. Various approaches are used to predict the regulatory edges. These methods can be classified to gene- or TF-centered approaches (Yang et al., 2016) or categorized as

#### 3.1 Construction of physical GRN

The edges in a physical network represent interactions between a TF and the specific CREs of the target genes it regulates (Schmitz et al., 2022). It is important to note that the physical interaction edges do not imply a functional alteration in gene expression. Instead, they represent a regulation potential that contributes to the complex transcription process.

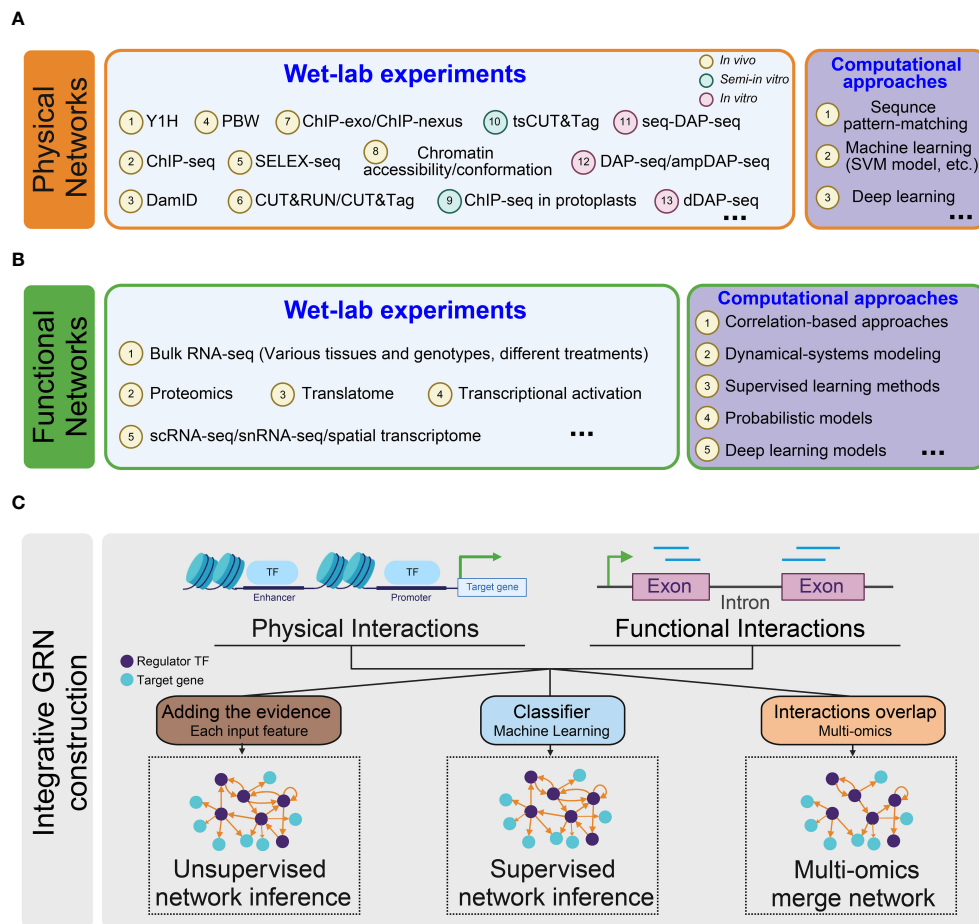


FIGURE 2

Overview of methodologies for constructing regulatory networks. (A) Inference of physical regulatory networks. Two types of methods are employed to construct physical networks: wet-lab experiments (light blue) and computational approaches (purple); (B) Inference of functional regulatory networks. Functional networks are inferred using two types of methods: wet-lab experiments (light blue) and computational approaches (purple); (C) Inference of integrative regulatory networks. Three types of methods are utilized to infer integrated gene regulatory networks: Unsupervised Network Inference (left): An integrative network is constructed by aggregating evidence from each input feature with equal weighting. Supervised Network Inference (middle): Input features are given to a classification model that predicts the presence or absence of a regulatory interaction for every TF-target pair. Multi-omics integration network (right): This approach identifies regulatory relationships using multi-omics data and merges them into a comprehensive network.

TFs bind to the genomic TF binding sites (TFBSs) through their DNA binding domains (DBDs). DBDs typically recognize short DNA motifs. Both experimental and computational approaches have been used to identify TFBS and DNA motifs recognized by specific DBDs. Large amount of TFBS and motif datasets have been collected and deposited in public databases, such as TRANSFAC (Matys et al., 2003), CIS-BP (Weirauch et al., 2014), JASPAR (Rauluseviciute et al., 2024), UniPROBE (Hume et al., 2015), PlantPAN (Chow et al., 2024), and ChIP-Hub (Fu et al., 2022).

### 3.1.1 Identification of TF-DNA interactions using experimental approaches

In addition to the yeast-one-hybrid (Y1H), chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) is a classical technique to identify TF-DNA interactions *in vivo* (Johnson et al., 2007; Gaudinier et al., 2011; Ferraz et al., 2021). New approaches have been developed to address the intrinsic limitations of ChIP-seq, such as its low resolution, low signal-to-noise ratio of detected peaks, and potential enrichment of non-targeted transcription factors (Worsley Hunt and Wasserman, 2014). For example, ChIP-exo and ChIP-nexus can improve the resolution of the detected peaks (He et al., 2015; Rossi et al., 2018). Techniques like CUT&RUN, CUT&Tag, and DamID can eliminate the need for crosslinking and significantly improve the sensitivity of detection (Meers et al., 2019; Alvarez et al., 2020; Tao et al., 2020). However, these methods are still costly and have even more technical complexity, limiting their applications.

Recently, new *in vivo* methods have been developed for large-scale experiments, utilizing tagged TFs transiently expressed in plant protoplasts. These modified ChIP-seq techniques can profile genome-wide TFBSs in an easier and relatively low-cost manner (Lee et al., 2017a; Tu et al., 2020). To further decrease the cost and improve detection sensitivity, Wu and coworkers introduced a transient and simplified CUT&Tag (tsCUT&Tag) method that involves the transient expression of tagged TFs in protoplasts combined with an improved CUT&Tag approach (Wu et al., 2022). This method promises to profile TFBSs more efficiently and cost-effectively across different plant species. Nevertheless, these protoplast-based methods restrict the obtained TF-DNA binding information to the specific tissue source of the protoplasts, such as leaves.

In contrast to *in vivo* methods, *in vitro* approaches eliminate the prerequisite for preparing antibodies specific to TFs or generating transgenic lines containing tagged TFs of interest. Therefore, they can be easily applied in a high-throughput manner (Ferraz et al., 2021). In protein binding microarrays (PBM) and systematic evolution of ligands by exponential enrichment-sequencing (SELEX-seq) methods (Berger and Bulyk, 2006; Smaczniak et al., 2017), TFs or TFs complexes, along with either immobilized DNA oligonucleotides or a random DNA library, are used in these tests. One limitation of PBM is that it may overlook longer TFBSs because it relies on short DNA oligos (10–12 base pairs). Similarly, DNA substrates in SELEX assays are not derived from genomic sequences and cannot be mapped to the genome. To address these limitations, DAP-seq is a recently developed method that utilizes fragmented

genomic DNA. In this way, DAP-seq captures more native genomic features, such as DNA methylation and the flanking sequences of core motifs (Bartlett et al., 2017). Nonetheless, DAP-seq method does not fully capture chromatin state information or the cofactors of TFs that can influence TF-DNA binding *in vivo*. To solve these issues, modified methods such as sequential DAP-seq (seq-DAP-seq) and double DAP-seq (dDAP-seq) techniques have been developed (Lai et al., 2020; Li et al., 2023a). In seq-DAP-seq, a sequential purification based on multiple tags was used. Lai and colleagues determined the genome-wide binding of the SEP3 homomeric complex using this method (Lai et al., 2020); dDAP-seq was used to elucidate the DNA binding and specificity of bZIP TF heterodimers and homodimers in *Arabidopsis*. This study demonstrates that heterodimerization of C/S1 family bZIP TFs expands their DNA binding preferences (Li et al., 2023a). To better reflect *in vivo* TF binding events, the DAP-seq data can be combined with accessible chromatin regions (ACRs) identified by ATAC-seq, DNase-seq, and MNase-seq. The results filtered by tissue- or cell-type-specific ACRs provide more accurate TFBSs by considering *in vivo* chromatin states.

### 3.1.2 Computational approaches used to predict the TF-DNA interactions

Taking advantage of the extensive experimentally determined TF-DNA interaction data, computational methods have been developed to make *de novo* prediction of TFBS for a given TF. Quantitative models of DNA motifs, such as the Position Weight Matrix (PWM), are required to depict the TF-DNA binding affinity and predicting new DNA binding sites (Jayaram et al., 2016). PWM-based motifs are often built using tools implemented in the HOMER (Heinz et al., 2010) and MEME Suite (Bailey et al., 2015) software collections. These motif discovery algorithms utilize a collection of TFBSs derived from ChIP-Seq, ATAC-seq data, or promoter analyses. Although PWMs provide a good approximation, this conventional model could be further improved by integrating sequence dependencies and DNA shape features (Lai et al., 2019).

Generally, two types of *in silico* approaches can be used to predict the TF binding to the genomic TFBSs: one relies on simple sequence pattern matching, and the other utilizes machine learning algorithms. The pattern matching-based algorithms follow the principle that candidate DNA binding sites possess sequence similarity with known DNA binding motifs of a TF. Several motif search algorithms, such as FIMO, MOODs, and PWMScan, are frequently adopted for this purpose (Grant et al., 2011; Korhonen et al., 2017; Ambrosini et al., 2018). Recent research indicates that pattern-matching-based methodology can be effectively applied across a diverse range of organisms (Puig et al., 2021; Chow et al., 2024). In plant species, PlantRegMap, which now incorporates PlantTFDB V5.0, is a major source for the inferred TF-DNA interactions. It now covers 165 species across the main lineages of green plants (Jin et al., 2017; Tian et al., 2020). However, due to the availability of known DNA motifs, most of the TF-DNA interactions have been identified in *Arabidopsis*.

Machine learning-based approaches establish predictive criteria by learning from documented TF-TFBS data using diverse computational strategies. For instance, Lee and coworkers introduced a SVM model incorporating various features of TFs and TFBSs, achieving approximately 82% prediction accuracy (Lee et al., 2017c). A recent study achieved a remarkable 99% accuracy in model prediction by integrating the chemical properties of TF proteins, along with the structural conformation and bonding capabilities of both TFs and DNA (Khamis et al., 2018). In plants, a SVM model was constructed to identify potential TFBSs for auxin response factor TFs in *Arabidopsis* (Cui et al., 2014). The TSPTFBS (v2.0) employed deep learning to model a total of 389 plant TFs with their binding sequences and achieved better performance than other standard methods (Cheng et al., 2023). Ruengsrachaiya and colleagues developed another machine-learning based predictor (Plant-DTI). This tool leverages a large number of experimental TF-TFBS interactions from plant species with a novel feature construction, resulting in a pronounced high predictive performance compared to other state-of-the-art methods (Ruengsrachaiya et al., 2022).

### 3.1.3 Connecting TF and target genes with chromatin accessibility and conformation data

Epigenetic modifications and chromatin states are essential factors in regulating gene expression. *In vivo*, most TFs bind to their target CREs within ACRs (Schmitz et al., 2022). Therefore, the identification of ACRs is an important aspect in the study of transcriptional regulation. Optimized genome-wide assays, such as DNase-seq (Boyle et al., 2008), MNase-seq (Mieczkowski et al., 2016), ATAC-seq (Buenrostro et al., 2013), and FAIRE-seq (Simon et al., 2012), have enabled the profiling of chromatin accessibility in numerous species and tissues. Currently, ATAC-seq has emerged as a prominent technique owing to its requirement of a reduced amount of nuclei input and the simplicity of its protocol (Buenrostro et al., 2013). Single-cell ATAC-seq (scATAC-seq) protocols have also been developed and optimized to allow the detection of open chromatin in individual plant cells (Buenrostro et al., 2015; Marand et al., 2021).

In addition to supporting and refining other regulatory networks (Duren et al., 2017), ACR datasets could be directly used in linking TFs to their target genes. This type of network inference pipeline consists of two main steps. Firstly, motif matcher algorithms, provided with TF binding motif data, are used to determine the interactions of TFs with accessible CREs. While scanning TF motifs in ACRs is the routine way, more and more advanced deep learning-based methods are employed to predict the TF binding sites directly from ATAC-seq data (Cazares et al., 2023). Then, these CREs are linked to genes based on a simple distance cutoff or a more refined assignment. These association relationships are combined to obtain “TF-CRE-gene” links and simplified to TF-gene pairs. The GRN inference based on ATAC-seq data can be accomplished with several software packages, such as ATAC2GRN (Pranzatelli et al., 2018), LISA (Qin et al., 2020), SPIDER (Sonawane et al., 2021), and MINI-AC (Manosalva Perez et al., 2024).

Although CREs bound by regulator TFs are often assigned to target genes based on closest genomic proximity, this simplistic approach may miss crucial distal interactions that have regulation effects. Accurately linking CREs to genes can be a challenge task, especially in large genomes that have many distantly located regulatory sequences (Ricci et al., 2019; Joly-Lopez et al., 2020). Chromatin is highly organized to form a three-dimensional (3D) structure. Techniques for measuring chromatin conformation, such as Hi-C and ChIA-PET, were used to capture the long-range chromatin interactions (Lieberman-Aiden et al., 2009; Ouyang et al., 2020). DNA conformation data has been successfully integrated with both ATAC-seq and RNA-seq data to construct GRNs (Jiang et al., 2022b). The characterization of gene regulatory systems based on 3D proximity can be achieved using methods such as DC3 (De-Convolution and Coupled-Clustering) (Zeng et al., 2019).

## 3.2 Inference of functional GRN

A functional regulatory network is characterized by TF-target edges that are supported by changes in the expression patterns of the target genes. These connections, whether direct or indirect, can reflect the functional impact of the regulator’s actions on their targets. From bulk RNA-seq data to single-cell transcriptomics, advanced inference methods have been developed, demonstrating enhanced accuracy and computational efficiency (Marbach et al., 2012a; Pratapa et al., 2020). Several approaches, which utilize time-series data or pseudo-temporal single-cell transcriptomics, were designed to gain more precise insights into the regulatory interactions between genes (Huynh-Thu and Geurts, 2018; Aubin-Frankowski and Vert, 2020).

Several statistical approaches are used for the transcriptome-based gene network analysis. The underlying principles of these methods include correlation, supervised learning, probabilistic models, dynamical-systems modeling, and deep learning (Li et al., 2015; Kim et al., 2023).

### 3.2.1 Correlation-based approaches

Co-expressed genes are believed to be functionally relevant or co-regulated. A regulatory link between TF and its target may be assumed by the co-expression pattern. The Pearson correlation coefficient (PCC), Spearman’s rank correlation coefficient (SCC), and mutual information (MI) coefficient are popular measures of gene’s co-expression patterns.

PCC is suitable for detecting linear correlations, whereas SCC is more robust to nonlinear relationships. Compared to linear correlation, nonlinear correlation is capable of detecting more complex relationships, which may better reflect the *in vivo* regulatory interactions (Zuin et al., 2022). The MI coefficient is a method based on information theory. It quantifies the interdependence between two variables and can detect nonlinear relationships (Song et al., 2012). However, co-expression analysis is unable to distinguish direct and indirect connections. There may be

two ways to solve this issue. One involves the computation of partial correlation coefficients among genes (de la Fuente et al., 2004); the other entails incorporating additional evidence from other data sources, including TF-DNA bindings and ATAC-seq.

### 3.2.2 Supervised learning methods

Supervised learning methods, such as linear regression, nonlinear regression, and tree-based approaches, are widely used for regulatory network construction.

The linear regression approach first collects expression data for a set of genes as the predictor variables and then regress on the expression levels of designated regulator TFs (response variable). The limitations of regression models include the risk of overfitting due to a large number of predictors (which is a common case in biological systems), and the challenges associated with high-dimensional data. These factors collectively impede the accurate inference of gene regulatory networks (Kim et al., 2023).

In contrast to linear regression, tree-based techniques like random forests have the ability to capture complex non-linear associations among genes (Huynh-Thu et al., 2010). These methods recursively divide the data into smaller subsets based on the predictor variables, creating a tree-like structure of decision rules. Each tree branch represents a distinct combination of predictor values, leading to a predicted value for the target gene at the leaf nodes. Notably, in the DREAM5 challenge (Marbach et al., 2012a), inference tools employing the random forest algorithm achieved the superior overall performance. However, these non-parametric approaches are often less interpretable than linear models. Additionally, they can be computationally intensive, especially when dealing with high-dimensional datasets.

### 3.2.3 Probabilistic models

Probabilistic models combine principles from probability theory and graph theory to construct networks. These methods capture the dependence between variables, such as transcription factors and their target genes, by modeling the presence and strength of regulatory relationships. Bayesian and Markov are two main types of probabilistic models.

In a Bayesian network, the target gene expression levels are assumed to follow a normal distribution conditioned on the expression levels of TF (Friedman, 2004). Bayesian networks are directed graphs that represent causal relationships between TFs and targets. However, they are unable to reflect feedback regulation relationships, as they do not have loops in the graph structure.

### 3.2.4 Dynamical-systems modeling

Dynamical systems-based approaches estimate the temporal expression patterns of genes. The regulatory influences of TFs, basal transcription, and inherent stochasticity can be modeled as parameters in differential equations (Hecker et al., 2009). Unlike regression and probabilistic-based approaches, dynamical-systems not only account for the diverse factors that regulate gene expression but also incorporate stochasticity. For example, the

observed expression variation among individual cells is biologically meaningful in single-cell RNA-seq data. Dictys method has been developed to utilize the influencing factors through an empirical linear stochastic differential equation (Wang et al., 2023a). It can capture changes in regulatory activity that are not solely dependent on gene expression levels, making it well-suited for studying continuous processes like cell differentiation (Wang et al., 2023a).

### 3.2.5 Deep learning models

Deep learning models, based on artificial neural networks, offer versatile architectures capable of performing various tasks (Min et al., 2017). Unlike other methods, deep learning models show increasingly improved performance as the size of the training dataset increases. Additionally, the feature extraction process is automatic, whereas other machine learning models require manual configuration.

Deep learning models excel in processing large datasets and approximating continuous relationships within the data, making them highly suitable for handling single-cell data to infer functional GRNs. A notable application is the use of autoencoders for dimension reduction and identifying potential regulatory relationships from various types of single-cell omics input data (Liu et al., 2023a). Additionally, many innovative approaches have emerged to utilize the matched scRNA-seq and scATAC-seq data (Ma et al., 2023a; Yuan and Duren, 2024). For example, Song and colleagues have introduced the multi-task-based MTLRank framework, which incorporates RNA velocity and scATAC-seq to obtain more accurate tissue-specific regulatory networks (Song et al., 2023). However, the application of these novel methods remains limited in plant species (Guo et al., 2024).

While deep learning models demonstrate their flexibility and ability to capture complex patterns, they often require large training datasets and substantial computational resources due to the vast number of parameters involved. Moreover, the models can be less interpretable than traditional models (Ma and Xu, 2022).

It is noteworthy that each model has its pros and cons (Marbach et al., 2012a). For example, correlation coefficient methods are more reliable for loop connections, whereas regression methods are suitable for linear regulatory relationships. Thus, the combination of multiple methods is expected to outperform individual methods (Vignes et al., 2011; Slawek and Arodz, 2013; Zhong et al., 2014).

## 3.3 Integrative GRN construction

In line with the concept of combining different methods in predicting functional GRN, combining physical and functional interactions datasets is also essential to construct comprehensive and high-confidence regulatory networks. The integration process can be achieved by simply using the ChIP-seq data for a TF along with the matched RNA-seq data in the mutant or by employing more advanced algorithms to merge the information from various multi-omics datasets.



### 3.3.1 Innovative approaches for aggregating TF-binding and gene expression datasets

The process of identifying direct and functional targets of a TF can be achieved by intersecting the TF-binding derived targets with differentially expressed genes identified from perturbations such as overexpression or knock-out of the TF. This method is considered state-of-the-art in TF target identification. However, it is important to note that these two evidence sources rarely converge on a common set of target genes. Despite being widely used as the gold standard, even the bound and differentially expressed genes may not be the validated functional targets (Kang et al., 2020).

To improve the prediction performance of TF-target relationships, a few advanced strategies have been proposed. Kang and coworkers introduced a method called Dual Threshold Optimization (DTO). This method improves the accuracy of identifying direct functional targets by combining data from TF binding sites and TF perturbation responses. The DTO method enhances the convergence of two data types by optimizing the significance thresholds for binding and responsive data (Kang et al., 2020). Morin and colleagues built upon two existing strategies (Tang et al., 2011; Wang et al., 2013) to create a framework to identify and rank TF-target interactions, and identified potential orthologous interactions between humans and mice. This workflow can be scaled to other TFs and offered experimental-level gene summaries evaluated against independent literature evidence (Morin et al., 2023).

### 3.3.2 Machine-learning based integration framework

To integrate more layers of physical and functional input data, more sophisticated machine-learning methods have been developed for integrative network inference (Mahood et al., 2020). The machine-learning-based methods can be grouped as unsupervised and supervised approaches.

The supervised approach utilizes a regression classifier, which is trained on known regulatory interactions to predict whether an edge (regulatory interaction) exists between TFs and target genes. In contrast, the unsupervised method averages the evidence across different feature-specific networks to generate a comprehensive regulatory network without requiring prior knowledge of regulatory interactions (Marbach et al., 2012b). Both the supervised and unsupervised integrative networks show high coverage. Recently, De Clercq and coworkers applied a supervised learning approach to integrate information about TF-binding, chromatin accessibility, and expression-based regulatory interactions in *Arabidopsis*. The resulting integrated GRN demonstrated high predictive power, facilitating the discovery of previously unidentified regulators (De Clercq et al., 2021).

## 4 Evaluation and downstream analyses of GRNs

After constructing a GRN, evaluating its accuracy and coverage is an essential task. And subsequent downstream analyses can be conducted to extract more biological insights.

## 4.1 Network evaluation

One should bear in mind that the connections in a GRN are hypothetical and require vigorous evaluation of their accuracy. Several common practices have been established to evaluate the biological relevance of the inferred connections (Li et al., 2015).

The most common evaluation method involves comparing the inferred GRN with a “gold standard” network, which is often derived from experimentally verified results, such as loss and gain of function experiments. These wet-lab approaches generate confident regulatory connections by observing the impact of a regulator’s expression changes on its target gene (Kim et al., 2023). When performing the comparison, one may calculate the average accuracy according to the detection ratio of the verified edges, which is probably flawed due to the sparsity of GRNs. In other words, an algorithm that always predicts the absence of edges could incorrectly achieve high accuracy. Thus, a better approach is to assess the proportion of correctly identified positives relative to all positives (sensitivity or recall) and the proportion of correctly identified positives out of all identified positives (precision or positive predictive value) (Huynh-Thu and Sanguinetti, 2019).

When experimentally validated “gold standard” or any well-accepted high-confidence networks are unavailable, alternative approaches for evaluating gene networks may include cross-validation tests and functional coherent module assessment. Cross-validation tests the accuracy of the reconstructed network by predicting gene functions based on the known functions of network neighbors. Additionally, high-quality networks are expected to exhibit coherent modules of interacting and co-regulated genes. The functional coherence of these modules can be evaluated through enrichment tests of gene function and probabilistic models to predict gene expression within the module (Li et al., 2015).

## 4.2 Downstream analysis

In addition to connecting TFs and their target genes, GRNs can provide further insights into gene functions and associated biological processes through downstream analysis.

### 4.2.1 Network topological analysis

GRNs often consist of large number of nodes and connections, which renders direct interpretation. Topological analysis has emerged as a useful method for examining the structural properties of these networks, such as node degree distribution, clustering coefficients, and community structures, to detect important patterns and anomalies within the network. In addition to uncovering the underlying structure of the graph, network topological analysis can also assist in identifying influential nodes or edges within the network. For instance, node centrality measures like degree centrality, betweenness centrality, and eigenvector centrality can highlight the most critical nodes in terms of their connectivity and impact on the network. Modularity is another important property of GRN (Segal et al., 2003). Genes within the

same module are often co-regulated and often share biological functions. Module detection helps identify sets of genes associated with specific biological processes. For example, Tu and colleagues partitioned a GRN of maize leaves into seven modules. Subsequent analyses using GO terms and MapMan revealed the enrichment of specific functions in each module (Tu et al., 2020).

#### 4.2.2 Comparative gene network analysis

Comparative analysis of GRNs can be used to compare different species, cell types, and treatment conditions. This approach provides more insight than directly comparing sequences or genes (Movahedi et al., 2011; Weston et al., 2011). During interspecific comparisons, it is important to conclusively define gene orthology and to ensure that comparable tissues are being examined.

Previous comparative GRN analysis methods involved pairwise subtraction of TF-gene interactions between GRNs (Thompson et al., 2015; Duren et al., 2021). However, due to the sparse and noisy nature of GRNs, a direct comparison of TF-gene interactions is not good enough. New strategies, such as topic modeling, have been employed to generate dense, low-dimensional representations that filter out the noise in the GRN and more robustly depict the differences in regulatory relationships (Lou et al., 2020).

#### 4.2.3 Prioritizing functional candidate regulators

Pinpointing the key regulatory TF in a network is of great interest in GRN downstream analyses. One approach to identifying these key TFs is to infer TF activities in a specific context using enrichment methods. These methods integrate gene expression with the topological information of GRNs, thereby extracting insights regarding the roles of TFs in particular biological contexts.

Commonly used methods for enrichment analysis include Gene Set Enrichment Analysis (GSEA) and Analysis of Upstream Regulators (AUCell) (Subramanian et al., 2005; Van de Sande et al., 2020). These techniques allow for a thorough analysis that integrates gene expression patterns with the structures of connections. For instance, Yuan and colleagues have utilized the AUCell enrichment method to discover high-activity TFs for each distinct cell type in a maize endosperm single-cell RNA-seq study (Yuan et al., 2024).

Moreover, the application of more sophisticated machine learning models has further advanced the prioritization of TFs. With the known-function genes as training data, these models are capable of identifying the TFs most significantly associated with specific biological processes. For example, NeuralNet algorithm was used to prioritize tassel branch number-related candidate genes (Wang et al., 2023b). Han and coworkers used a similar approach to generate a prediction model based on an integrative map, and predicted which genes are associated with the flowering time pathway (Han et al., 2023).

## 5 Recent advances in regulatory network studies of crop species

In network-related literatures, some focus on developing new inference methods or serving as database resources (Table 1); others

are dedicated to solving specific biological questions. Many studies with advanced concepts have been conducted in the model plant *Arabidopsis* (Table 2).

For example, Hickman and colleagues conducted a time-series experiment to study the regulation of JA response in *Arabidopsis*. They used RNA-seq data from 14-time points on MeJA (methylated ester of JA) treated leaf and constructed a dynamic model of the JA GRN. This study offers significant advances in our understanding of how plants dynamically regulate the JA signaling pathway in response to environmental cues and lays a foundation for further investigating the complex transcriptional programs underlying plant stress responses and developmental processes (Hickman et al., 2017). Zender and coworkers combined time-series transcriptome, proteome, and phosphoproteome data to reconstruct GRNs, predict new components involved in the JA signaling pathway, and validate these new genes through genetic mutants. This work demonstrates the power of integrative multi-omics approach to provide fundamental biological insights into plant hormone responses (Zander et al., 2020). De Clercq and colleagues have combined networks based on DNA motifs, open chromatin, transcription factor (TF) binding, and expression-based interactions through a supervised learning approach. The integrated GRN outperforms the individual input networks in predicting known regulatory interactions. They also experimentally validated many TFs involved in reactive oxygen species (ROS) stress regulation, including 13 novel ROS regulators (De Clercq et al., 2021).

Researchers also construct many regulatory networks in crop species, particularly in cereals such as maize and wheat. We have endeavored to summarize these works with a focus on presenting the cutting-edge findings rather than aiming for comprehensiveness. We highlight a selection of literatures from the last decade (Table 3).

### 5.1 GCN studies in crops

Early-stage studies often relied solely on bulk gene expression data, typically obtained from specific plant tissues or organs, to construct functional GCNs.

For example, Pfeifer and colleagues analyzed gene expression in developing wheat grains and constructed a co-expression network comprising 25 modules. These modules displayed unique spatiotemporal characteristics that can be distinguished based on grain cell types or developmental stages (Pfeifer et al., 2014). To provide insights into the coordination of individual homoeologs underlying various traits in wheat, the coexpression networks were constructed from nonstress tissue-specific and stress-related RNA-seq samples. These networks highlight the extensive coordination of homoeologs throughout development and in response to various stresses and offer a platform to identify candidate genes for agronomic traits (Ramírez-González et al., 2018). Huang and coworkers evaluated various parameters for data normalization and different inference methods for constructing a large GCN in maize using RNA-Seq data. The analysis revealed that increasing sample size positively impacts network performance, emphasizing

**TABLE 1** Database resources of regulatory network from the past decade.

Species	Database Name	References
Multiple species	ATTED-II (v11)	(Obayashi et al., 2022)
Multiple species	ATTED-II	(Obayashi et al., 2018)
Multiple species	PlantTFDB (v3.0)	(Jin et al., 2014)
Multiple species	PlantTFDB (v4.0)	(Jin et al., 2017)
Multiple species	PlantRegMap	(Tian et al., 2020)
Multiple species	PlantPAN (v2.0)	(Chow et al., 2016)
Multiple species	PlantPAN (v3.0)	(Chow et al., 2019)
Multiple species	PlantPAN (v4.0)	(Chow et al., 2024)
Multiple species	ChIP-Hub	(Fu et al., 2022)
Multiple species	KnockTF	(Feng et al., 2024)
Multiple species	Plant-DTI	(Ruengsrachaiya et al., 2022)
Multiple species	PlaPPISite	(Yang et al., 2020)
<i>Arabidopsis</i> / Animal species	UniBind	(Puig et al., 2021)
<i>Arabidopsis</i> /Maize/Rice	ConnecTF	(Brooks et al., 2021)
Soybean	SoyNet	(Kim et al., 2017a)
Tomato	TomatoNet	(Kim et al., 2017b)
<i>Arabidopsis</i>	TF2Network	(Kulkarni et al., 2018)
<i>Arabidopsis</i>	Cistrome	(O'Malley et al., 2016)
<i>Arabidopsis</i>	AraNet (v2)	(Lee et al., 2015b)
<i>Arabidopsis</i>	AraPPINet	(Zhao et al., 2019)
<i>Arabidopsis</i>	AGRIS	(Yilmaz et al., 2011)
Rice	RicePPINet	(Liu et al., 2017)
Rice	RiceENCODE	(Xie et al., 2021)
Rice	NetREx	(Sircar et al., 2022)
Rice	RiceTFtarget	(Zhang et al., 2023a)
Rice	RiceNet (v2)	(Lee et al., 2015a)
Maize	MaizeNetome	(Feng et al., 2023)
Maize	CORNET (v2.0)	(De Bodt et al., 2012)
Maize	MaizeNet	(Lee et al., 2019)
Wheat	WheatNet	(Lee et al., 2017b)
Wheat	WheatOmics	(Ma et al., 2021)
Wheat	wGRN	(Chen et al., 2023b)
Wheat	Wheat-RegNet	(Tang et al., 2023)
Wheat	WheatCENet	(Li et al., 2023b)

the importance of sample size for the construction of accurate GCNs (Huang et al., 2017). To extend the knowledge of salt response in soybean, Hu and colleagues clustered differentially expressed genes between a salt-tolerant and a salt-hypersensitive

cultivar. They constructed undirected networks representing their co-expression patterns based on Pearson's correlation coefficients. The network analysis unveiled several candidate pathways critical in salt responses, including phytohormone signaling, oxidoreduction, phenylpropanoid biosynthesis, and others (Hu et al., 2022).

While GCNs cannot directly define the regulatory relationships between TFs and their downstream targets, the connectivity results are widely used to refine the findings from genome-wide association studies (GWAS) and quantitative trait locus (QTL) analyses. This integration enables the effective prediction of novel candidate genes. For example, a weighted GCN analysis was used to identify connected genes associated with Fusarium head blight (FHB) resistance and pinpointed candidate hub genes within the interval of three previously reported FHB resistance QTL in wheat (Sari et al., 2019). Yao and coworkers combined GWAS and co-expression network analyses to uncover candidate genes involved in the accumulation of oleic acid content in rapeseed (Yao et al., 2020). GCN analysis and genome-wide association studies (GWAS) were combined to elucidate the regulatory pathways and identify candidate genes responsible for pre-harvest sprouting and seed dormancy traits in maize (Ma et al., 2023b).

## 5.2 Networks based on TF-DNA binding in crops

Establishing direct physical interactions between TF and DNA has been a major research focus. In addition to TFBSs obtained from experimental techniques such as ChIP-seq, many TFBSs have been predicted by computational algorithms. For example, Yu et al. collected transcriptomics data from developing maize leaves and used co-expression data along with enrichment analysis to predict overrepresented motifs in the promoter sequences and the potential TFBSs of key TFs (Yu et al., 2015).

A few databases and newly developed inference methods have significantly expanded the available information on TF-DNA binding interactions. PlantPAN, which has collected a comprehensive set of public ChIP-seq datasets, is a valuable resource for plant TF-TFBS interactions. It offers the most complete plant PWMs for analyzing TFBSs and effective tools for predicting TFBSs in conserved regions of a given promoter. The latest version, PlantPAN 4.0, provides a non-redundant set of 3,428 matrices for 18,305 TFs of 115 plant species (Chow et al., 2024). Another valuable resource is ChIP-Hub, a comprehensive and standardized platform for exploring the regulome of plants. It collects over 10,000 datasets from 41 plant species and processes them based on ENCODE standards. As an application example, an extensive survey was performed to examine the co-associations among various regulators, enabling the construction of a hierarchical regulatory network spanning a broad developmental context (Fu et al., 2022).

Meanwhile, wet-lab approaches persist in being actively employed to extend experimental TF-DNA binding in plants. In wheat, Zhang and colleagues have successfully obtained high-quality DNA binding profiles for 53 environmentally responsive TFs using DAP-seq. Interestingly, the study found that 85% of the *in vitro* identified TFBSs were located within transposable elements and

TABLE 2 Selected network-related studies in *Arabidopsis* from the past decade.

Species	Research Objective	Brief Description	References
<i>Arabidopsis</i>	Methodology	MICRAT	(Yang et al., 2018)
<i>Arabidopsis</i>	Biology question	PPI network	(Nietzsche et al., 2016)
<i>Arabidopsis</i>	Biology question	GRN of root	(Santuari et al., 2016)
<i>Arabidopsis</i>	Biology question	GRN of root stem cell	(de Luis Balaguer et al., 2017)
<i>Arabidopsis</i>	Biology question	JA signal pathway (GRN)	(Hickman et al., 2017)
<i>Arabidopsis</i>	Methodology	CoReg package	(Song and Li, 2023)
<i>Arabidopsis</i>	Biology question	Metabolic Pathways	(Wisecaver et al., 2017)
<i>Arabidopsis</i>	Biology question	Flower development (GRN)	(Chen et al., 2018)
<i>Arabidopsis</i>	Methodology	Protein binding microarrays (PBM)	(Franco-Zorrilla et al., 2014)
<i>Arabidopsis</i>	Biology question	Nitrogen metabolism	(Gaudinier et al., 2018)
<i>Arabidopsis</i>	Methodology	Chromatin accessibility	(Sullivan et al., 2014)
<i>Arabidopsis</i>	Biology question	Single and combined stresses	(Barah et al., 2016)
<i>Arabidopsis</i>	Biology question	Nitrogen signaling	(Varala et al., 2018)
<i>Arabidopsis</i>	Biology question	Response to elevated CO <sub>2</sub>	(Cassan et al., 2023)
<i>Arabidopsis</i>	Biology question	NLP7 regulon	(Alvarez et al., 2020)
<i>Arabidopsis</i>	Biology question	Nitrogen signaling	(Brooks et al., 2019)
<i>Arabidopsis</i>	Methodology	EXPLICIT	(Geng et al., 2021)
<i>Arabidopsis</i>	Methodology	EXPLICIT-Kinase	(Peng et al., 2022)
<i>Arabidopsis</i> /Rice/Maize	Methodology	TSPTFBS 2.0	(Cheng et al., 2023)
<i>Arabidopsis</i>	Methodology	ConsReg	(Song et al., 2019)
<i>Arabidopsis</i> /Rice	Methodology	Comparative analysis of GRN	(Wang et al., 2020b)
<i>Arabidopsis</i>	Biology question	Response to JA	(Zander et al., 2020)
<i>Arabidopsis</i>	Biology question	ROS signaling	(De Clercq et al., 2021)
<i>Arabidopsis</i>	Methodology	Annotation of unknown gene	(Depuydt and Vandepoele, 2021)
<i>Arabidopsis</i>	Biology question	scATAC-seq of root	(Dorrity et al., 2021)
<i>Arabidopsis</i> /Rice/Maize	Methodology	MINI-EX	(Ferrari et al., 2022)
<i>Arabidopsis</i> /Rice/Maize/ Tomato	Methodology	MINI-EX V2.0	(Jasper et al., 2023)
<i>Arabidopsis</i> /maize	Methodology	MINI-AC	(Manosalva Perez et al., 2024)
<i>Arabidopsis</i>	Biology question	scRNA-seq of BR root	(Nolan et al., 2023)
<i>Arabidopsis</i> /Maize	Biology question	Nitrogen responsive GRN	(Cheng et al., 2021)

associated with regulatory sequences specific to the wheat lineage (Zhang et al., 2021). In a subsequent study by the same group, genomic binding profiles were generated for a larger set of TFs, enabling the assembly of a wheat GRN encompassing connections among 189 TFs and 3,714,431 regulatory elements (Zhang et al., 2022b). These results provide valuable insights into the transcriptional regulatory mechanisms in wheat. Several remarkable advances were also made in maize. Ricci and coworkers performed DAP-seq on 32 TFs, indicating that the distal accessible chromatin regions were enriched for TFBS (Ricci et al., 2019). Additionally, interaction maps were generated for 14 maize TFs from the ARF

family, revealing both specific and redundant binding events of ARF TFs (Galli et al., 2018). Furthermore, 104 maize TFBS datasets were yielded by ChIP-seq with transient expressed proteins to construct the leaf regulatory network (Tu et al., 2020).

### 5.3 Inference of GRN using expression data in crops

GRNs derived from gene expression profiles are not limited by the availability of TF-DNA binding data and are widely used in various

TABLE 3 Selected network-related studies in crops from the past decade.

Species	Research Objective	Descriptions	References
Rice	Biology question	Pollen development	(Lin et al., 2017)
Rice	Biology question	Response to Cadmium stress	(Wang et al., 2020a)
Rice	Methodology	Improvement of PBM	(Kim et al., 2021a)
Rice	Biology question	Gene editing promoter of <i>IPA1</i>	(Song et al., 2022)
Rice	Biology question	Phosphate starvation response	(Shi et al., 2021)
Rice	Biology question	Biotic stress response	(Lee et al., 2011)
Rice	Biology question	scRNA-seq GRN	(Xie et al., 2020)
Rice	Biology question	Agronomic traits	(Zhang et al., 2022a)
Rice	Biology question	Low temperature response	(Wang et al., 2022a)
Maize	Biology question	Leaf development	(Yu et al., 2015)
Maize	Biology question	Leaf development	(Tu et al., 2020)
Maize	Biology question	Seed development	(Zhan et al., 2015)
Maize	Biology question	Seed development	(Liu et al., 2016)
Maize	Biology question	Seed development	(Yi et al., 2019)
Maize	Biology question	Seed development	(Xiong et al., 2017)
Maize	Biology question	Seed development	(He et al., 2024)
Maize	Biology question	Inositol phosphate metabolism	(Zhang et al., 2016b)
Maize	Biology question	Developmental atlas	(Walley et al., 2016)
Maize	Biology question	Gene network of grey leaf spot	(Christie et al., 2017)
Maize	Methodology	Optimize GCN construction	(Huang et al., 2017)
Maize	Biology question	Phenolic compound biosynthesis	(Yang et al., 2017)
Maize	Biology question	TFBS of ARF family	(Galli et al., 2018)
Maize	Biology question	Tissue-specific GRN	(Huang et al., 2018)
Maize	Biology question	Inflorescence development	(Parvathaneni et al., 2020)
Maize	Biology question	Meta GRNs using RNA-seq data	(Zhou et al., 2020)
Maize	Biology question	scATAC-seq of 6 tissues	(Marand et al., 2021)
Maize	Biology question	scRNA-seq of ear	(Xu et al., 2021)
Maize	Biology question	scRNA-seq of leaf	(Tao et al., 2022)
Maize	Biology question	Spatial transcriptomics of seed	(Fu et al., 2023)
Maize	Biology question	scRNA-seq of endosperm	(Yuan et al., 2024)
Maize	Biology question	Multi-omics integrative network	(Han et al., 2023)
Maize	Biology question	Early shade avoidance response	(Wang et al., 2016)
Maize	Biology question	Translatome-transcriptome GRN	(Zhu et al., 2023b)
Maize	Biology question	Prioritizing Metabolic GRN	(Gomez-Cano et al., 2024)
Maize	Biology question	Lipid metabolism	(de Abreu et al., 2018)
Maize	Biology question	UV-B stress response	(Gupta et al., 2018)
Maize	Biology question	Seed dormancy	(Ma et al., 2023b)
Maize	Biology question	Development, nutrients utilization, metabolism, and stress response	(Ma et al., 2017)

(Continued)

TABLE 3 Continued

Species	Research Objective	Descriptions	References
Maize	Biology question	Responses to Puccinia sorghi	(Kim et al., 2021b)
Maize	Biology question	Bundle sheath and mesophyll cells network	(Dai et al., 2022)
Maize	Biology question	WGCNA of bundle sheath and mesophyll cells	(Tao and Zhang, 2022)
Maize	Methodology	ChIA-PET	(Peng et al., 2019)
Wheat	Biology question	Grain transcriptome	(Pfeifer et al., 2014)
Wheat	Biology question	The transcriptional landscape of polyploid wheat	(Ramírez-González et al., 2018)
Wheat	Biology question	Regulating Senescence	(Borrill et al., 2019)
Wheat	Biology question	Fusarium head blight resistance	(Sari et al., 2019)
Wheat	Biology question	Embryogenesis and grain development	(Xiang et al., 2019)
Wheat	Biology question	Chloroplast biogenesis	(Loudya et al., 2021)
Wheat	Biology question	Evolutionary rewiring of the wheat transcriptional regulatory network	(Zhang et al., 2021)
Wheat	Biology question	Construction of GRN with DAP-seq data	(Zhang et al., 2022b)
Wheat	Biology question	Spike architecture	(Wang et al., 2017)
Wheat	Biology question	Biologically-Relevant	(Harrington et al., 2020)
Wheat	Biology question	Integrate gene regulatory network and genetic variation	(Ai et al., 2024)
Wheat	Biology question	Spike development	(Lin et al., 2024)
Wheat	Biology question	Regeneration	(Liu et al., 2023b)
Wheat	Biology question	scrRNA-seq of root	(Zhang et al., 2023b)
Sorghum	Biology question	Bioenergy stems	(Fu et al., 2024)
Cotton	Biology question	Oil accumulation	(Ma et al., 2024)
Cotton	Biology question	Low light intensity	(Zhao et al., 2024)
Cotton	Biology question	Seed yield	(Zhao et al., 2023)
Rapeseed	Biology question	Oleic acid content	(Yao et al., 2020)
Soybean	Biology question	Salt stress	(Yang et al., 2019)
Tomato	Biology question	Drought-responsive	(Chowdhury et al., 2023)
Sweet potato	Biology question	Chlorogenic acid biosynthesis	(Xu et al., 2022a)
Banana	Biology question	Fruit ripening	(Kuang et al., 2021)
Multiple species	Methodology	cisDyNet software	(Zhu et al., 2023a)

biological contexts. A GRN was inferred by modeling 78 maize seed transcriptome to identify key genes involved in seed development. The network analysis unraveled highly interwoven communities and identified key genes and regulatory modules associated with nutrient transport and imprinting patterns, which are crucial for maize seed development (Xiong et al., 2017). Utilizing the GENIE3 software package with a number of RNA-Seq data, Huang and colleagues constructed four tissue-specific GRNs in maize. They further predicted key TFs for each specific tissue (Huang et al., 2018).

Zhou and colleagues present a standardized pipeline using machine learning algorithms along with transcriptomic data to predict GRNs (Zhou et al., 2020). They analyzed a large collection of transcriptome datasets, resulting in 45 GRNs. The networks exhibited significant enrichment for biologically relevant

interactions, with each GRN capturing diverse biological processes. This uniform pipeline can be applied to other species with available expression data (Zhou et al., 2020). To comprehensively elucidate the chloroplast biogenesis process, Loudya and colleagues present a biologically informed GRN. The network prediction suggests that the regulators of chloroplast genes are differentially involved across various leaf developmental stages in wheat (Loudya et al., 2021).

Similar to GCN, GRN can also be combined with QTL and GWAS results to predict candidate genes for specific traits. For example, Zhao and colleagues designed an integrative analysis combining eQTL, GWAS, and GRN to characterize the genetic basis of cotton yield. Several high-ranking causal genes identified from the GRN were validated for their functional impacts on cotton seed development (Zhao et al., 2023).

## 5.4 Integrative network construction with multi-omics data in crops

Genomics and functional genomics studies on rice have been at the forefront among crop plants. Rice also serves as a leading model in integration network studies. The RiceNet (v2) web resource, launched in 2014, provides an integrative network for rice. This network combined co-functional links based on genomic context similarity, connections inferred from co-expression patterns, and protein-protein interactions. Its utility in prioritizing candidate genes involved in rice biotic stress responses has been demonstrated (Lee et al., 2015a). Another significant pioneering study created a comprehensive developmental atlas of maize with multi-omics data. Integrative GRNs were constructed based on mRNA, protein, and phosphor-protein data, resulting in improved predictive power. This work enhanced our understanding of the complex regulatory mechanisms in maize (Walley et al., 2016).

The integration of multi-omics data has become increasingly prevalent in studies using network-based approaches. Han and colleagues have successfully constructed a large-scale PPI network in maize. An integrated map was constructed incorporating data from four different layers: three-dimensional genomics, transcriptomics, proteomics, and protein-protein interactions. Leveraging this multi-omics network and machine learning-based prediction approaches, novel candidate key genes involved in various regulatory pathways, such as flowering time, have been predicted and genetically validated (Han et al., 2023). Gomez-Cano and coworkers analyzed ~4.6M interactions, including co-expression networks, TF-DNA interaction experiments, and expression quantitative trait loci (eQTL) to construct GRNs and pinpointed key regulators associated with hormone, metabolic, and developmental processes (Gomez-Cano et al., 2024). Additionally, several studies integrate a large amount of omics data, including both physical interactions and functional regulation relationships, in wheat (Chen et al., 2023; Tang et al., 2023). Similar integrated analysis has also been conducted in cotton (Zhao et al., 2024).

The integration of network and genetic mapping data, such as GWAS, further enhances the predictive power for identifying significant genes. Lin and colleagues thoroughly examined the transcriptome and epigenome profiles of the developing spike in an elite wheat cultivar. Through the integration of regulatory networks with GWAS, key genes affecting the spike architecture were pinpointed (Lin et al., 2024).

## 5.5 Regulatory networks at a high spatial resolution in crops

Regulatory networks relying on bulk data have several limitations. These models typically only capture generalized connection patterns, which obscure distinct regulatory interactions unique to certain cell types. Furthermore, bulk data often fails to differentiate the cellular states, which can significantly impact gene regulation. In contrast, approaches such as microdissection and single-cell technologies enable the discovery of regulatory networks at greater spatial resolution.

Zhan and coworkers used laser-capture microdissection with RNA-Seq to profile gene expression in each dissected cell compartment of the maize kernel (Zhan et al., 2015). They constructed an unbiased GCN and detected sub-network modules containing genes predominantly expressed in a single compartment or ubiquitously expressed across multiple compartments. These results offer a high-resolution gene expression atlas of maize kernel and contribute to uncovering regulatory interactions associated with the differentiation of major endosperm cell types (Zhan et al., 2015).

With the advent of single-cell omics, Marand and colleagues generated a cis-regulatory atlas using single-cell ATAC-seq in maize. They profiled over 72,000 nuclei across six maize organs and identified TFs coordinated with chromatin interactions by analyzing patterns of co-accessible CREs. This comprehensive cis-regulatory atlas at single-cell resolution is a valuable resource to study the gene regulation in maize (Marand et al., 2021). The researchers from Vandepoele's group have developed computational methods named MINI-EX and MINI-AC to explore cell type-specific regulatory interactions. MINI-EX utilizes expression-based GRNs derived from single-cell RNA-seq data and TF binding motifs to predict cell type-specific regulons. On the other hand, MINI-AC combines accessible chromatin (AC) data from either bulk or single-cell experiments with TF binding motifs to construct GRNs (Ferrari et al., 2022; Manosalva Perez et al., 2024). The application of MINI-EX has successfully identified regulons (groups of genes co-regulated by a shared TF) across major cell types in *Arabidopsis*, rice, and maize. Moreover, this method effectively prioritized established key regulons based on their network characteristics, such as connectivity and centrality, and unveiled several previously unidentified transcriptional regulators (Ferrari et al., 2022). Similarly, MINI-AC has also demonstrated superior performance compared to other techniques in accurately identifying TFBS. Maize has a complex genome and abundant distal AC regions. MINI-AC successfully inferred leaf GRNs containing experimentally confirmed interactions between TFs and target genes from both proximal and distal regions in maize. It is also a robust tool for pinpointing both known and novel candidate regulators (Manosalva Perez et al., 2024).

Recently, Yuan and coworkers focused on the differentiation stage of maize endosperm. They performed single-cell RNA-seq combined with TFBS profiling using ampDAP-seq to construct a high-confidence GRN and identified key regulators in five distinct cell types (Yuan et al., 2024). Fu and colleagues utilized the endosperm spatial transcriptome data during the grain-filling stage. They successfully predicted and identified the function of the candidate sucrose transporter genes (*SUTs*) in endosperm transfer cells facilitated by GCN analysis (Fu et al., 2023).

## 5.6 Utilizing clues from networks to answer biological questions in crops

Unlike traditional forward genetics, a new research paradigm is emerging for gene function studies, wherein candidate genes are

determined through hints from networks. Y1H is an easy approach used to identify the direct binding between TFs and the promoters of their targets. These direct regulatory relationships have great value in guiding the selection of key regulators for functional characterization.

For instance, Gaudinier and colleagues used enhanced Y1H assays to screen for *Arabidopsis* TFs binding to the promoters of genes associated with nitrogen metabolism and signaling, resulting in a network comprising 1,660 interactions. This network unveiled a hierarchical regulation of these TFs. Mutants of 17 prioritized key TFs exhibited significant alterations in at least one root architecture trait. The identification of regulatory TFs in the nitrogen-regulatory framework holds promise for enhancing agricultural productivity (Gaudinier et al., 2018). Similarly, Shi and coworkers uncovered TFs that regulate genes related to mycorrhizal symbiosis using Y1H. They screened more than 1500 rice TFs for binding to 51 selected promoters, and constructed a highly interconnected network. Interestingly, many of the TF in this network are involved in the conserved P-sensing pathway. With functional analyses of selected genes, this study elucidates the extensive regulation of mycorrhizal symbiosis by both endogenous and exogenous signals (Shi et al., 2021).

Ji and colleagues constructed a co-expression network to identify regulatory factors during the grain-filling stage of maize endosperm, and identified hundreds of candidate TFs using 32 storage reserve-related genes as guides. In addition to known regulators of storage proteins and starch, the study uncovered novel TFs, such as *GRAS11*, involved in endosperm development. They further characterized the function of *GRAS11* through detailed functional analysis (Ji et al., 2022). High-temporal-resolution RNA sequencing was conducted on the basal and upper regions of maize kernels. Weighted gene co-expression network analyses were performed, identifying numerous hub regulators that are worthy of subsequent functional characterization (He et al., 2024).

Collectively, these studies have provided significant insights into transcriptional regulation programs and rich data resources. GCN analyses have identified modules and candidate genes associated with various traits. Experimental determination of TFBSs, aided by computational predictions, has enabled the construction of regulatory networks, revealing novel regulators. Integration of multi-omics data has improved the predictive power of GRNs. High-resolution spatial techniques have uncovered cell-type-specific regulatory interactions, providing a more nuanced understanding of gene regulation. Overall, the advancements in regulatory network studies of crop species have substantially enhanced our understanding of the complex transcriptional programs governing plant growth, development, and responses to biotic and abiotic stresses.

## 6 Challenges and future perspectives

The precise manipulation of gene expression can be used to breed crops with desirable traits. The inference and analysis of regulatory networks will assist in crop improvement efforts. Despite the significant breakthroughs in regulatory network studies in

recent years, there is still potential for enhancing the confidence of the inferred interactions.

### 6.1 Network validation is a complicated task

The validation of regulatory networks is crucial to ensure that these networks accurately reflect the biological processes of interest. GRN evaluation commonly requires a thorough comparison of predicted interactions with the “gold standard” derived from wet-lab experiments, such as independently generated TF-DNA binding data and perturbation tests on the regulator TFs, as discussed in **Section 4.1**. However, the experimental “gold standard” is often unavailable or inadequate. Noteworthy, even if TF binding to a target is confirmed by *in vivo* ChIP assay, it does not necessarily imply that this TF can activate or repress the target gene. An alternative approach to test whether a TF regulates a particular gene is to perturb the TF expression and check how this perturbation affects the expression of target genes. While this approach shows promise due to its inherent causality, perturbation experiments are time consuming and costly. Noteworthy, they are likely not to work well as hindered by the widely existed compensatory mechanisms in crop plants. Due to different validation methods have their own limitations, utilizing diverse assessment strategies to evaluate a given GRN may be a smart way.

### 6.2 Increasing spatiotemporal resolution of networks

Regulation of gene expression is a dynamic process. High-temporal resolution studies have revealed fluctuations in gene expression levels in maize kernels within a small time window (Yi et al., 2019). From the perspective of network inference, there may be lack of expression correlation between target genes and their regulatory TFs because of the temporal lag between TF binding and the accumulation of mRNA transcripts. Thus, improving the spatiotemporal resolution of gene expression profiles and TF-DNA binding data is imperative for network construction.

Although new technologies have been developed to more efficiently acquire multi-omics data from a single plant cell (Xu et al., 2022b), the primary challenge arises from technical difficulties in the preparation of high-quality protoplasts in plants. Recently, to address challenges in protoplasting experiments, several optimized enzymatic cell wall digestion protocols have been developed for various species (Ye et al., 2022; Wang et al., 2022b; Chen et al., 2023a). Wang and colleagues introduced a new method that involves two consecutive digestion processes with different enzymatic buffers, significantly enhancing the efficiency and viability of protoplast preparation across diverse plant tissues (Wang et al., 2022b). However, the conversion of cells into protoplasts is still not feasible for many types of plant tissues. Alternatively, recent single-nucleus techniques offer broader applicability across different tissues. Nevertheless, it's important to note that nuclear RNA and cytoplasmic RNA should not be considered equivalent.



## 6.3 Hurdles in linking GRN to agronomic traits

There is still a large gap between the knowledge derived from GRNs and their manifestation in agronomic traits. Firstly, GRNs involve a complex interplay among thousands of genes and TFs that underlie various biological processes. Deciphering how perturbations in these regulatory networks impact gene expression remains a challenge. Secondly, the relationship between gene expression levels and traits is nonlinear and polygenic. Therefore, predicting observable traits from changes in gene expression, especially considering the influence of environmental factors, is difficult.

At the current stage, it is feasible to modulate specific metabolic pathways based on network information. The activity of one or a few TFs can regulate multiple steps of metabolic pathways. Thus, manipulating the expression of TFs probably has a greater impact on metabolism pathways than modifying cis-regulatory elements of enzyme-coding genes. For example, flavonoids are considered valuable compounds in plant metabolic engineering. Increasing flavonoid levels can be achieved by manipulating their transcription regulatory elements, resulting in the development of plants with high anthocyanin content (Jiang et al., 2023). For the enhancement of oil production, *WRINKLED1* is a conserved transcription factor involved in the regulation of fatty acid biosynthesis in diverse angiosperms. Transgenic plants that overexpress the *WRINKLED1* gene show promising outcomes in increasing the oil content of maize, soybean, and rice. As *WRINKLED1* also modulates targets that affect plant growth and development. It is important to consider the shared regulatory network when utilizing it to engineer plant oil production (Yang et al., 2022).

However, achieving the modulation of complex traits, such as yield and quality, which are determined by multiple factors, remains challenging. The regulatory mechanisms that directly impact these processes have not been thoroughly characterized. And these complex traits are often influenced by environmental factors and are sensitive to the interplay between genotype and environment.

## 6.4 Future directions

In response to current challenges, there are several aspects in future network-related research that need to be strengthened.

Firstly, integrating more omics data can enhance the predictive power of networks by merging diverse complementary information. Secondly, improving spatiotemporal resolution relies on the development of more sensitive, convenient, and cost-effective technologies. Lastly, the application of deep learning models, which can better integrate massive amounts of data and extract reliable and useful information from them, provides an opportunity to construct more accurate GRNs.

## Author contributions

QH: Writing – original draft, Writing – review & editing. RS: Writing – review & editing. ZM: Writing – review & editing, Writing – original draft.

## Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. The work is supported by the Frontiers Science Center for Molecular Design Breeding (2022TC146) and the 2115 Talent Development Program of China Agricultural University.

## Conflict of interest

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

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Ai, G., He, C., Bi, S., Zhou, Z., Liu, A., Hu, X., et al. (2024). Dissecting the molecular basis of spike traits by integrating gene regulatory network and genetic variation in wheat. *Plant Commun.* 5, 100879. doi: 10.1016/j.xplc.2024.100879
- Aibar, S., González-Blas, C. B., Moerman, T., Huynh-Thu, V. A., Imrichova, H., Hulselmans, G., et al. (2017). SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* 14, 1083–1086. doi: 10.1038/nmeth.4463
- Alanis-Lobato, G., Bartlett, T. E., Huang, Q., Simon, C. S., McCarthy, A., Elder, K., et al. (2024). MICA: a multi-omics method to predict gene regulatory networks in early human embryos. *Life Sci. Alliance* 7, e202302415. doi: 10.26508/lsa.202302415
- Altmann, M., Altmann, S., Rodriguez, P. A., Weller, B., Elorduy Vergara, L., Palme, J., et al. (2020). Extensive signal integration by the phytohormone protein network. *Nature* 583, 271–276. doi: 10.1038/s41586-020-2460-0
- Alvarez, J. M., Schinke, A. L., Brooks, M. D., Pasquino, A., Leonelli, L., Varala, K., et al. (2020). Transient genome-wide interactions of the master transcription factor NLP7 initiate a rapid nitrogen-response cascade. *Nat. Commun.* 11, 1157. doi: 10.1038/s41467-020-14979-6
- Ambrosini, G., Groux, R., and Bucher, P. (2018). PWMScan: a fast tool for scanning entire genomes with a position-specific weight matrix. *Bioinformatics* 34, 2483–2484. doi: 10.1093/bioinformatics/bty127
- Aubin-Frankowski, P. C., and Vert, J. P. (2020). Gene regulation inference from single-cell RNA-seq data with linear differential equations and velocity inference. *Bioinformatics* 36, 4774–4780. doi: 10.1093/bioinformatics/btaa576
- Badia-i-Mompel, P., Wessels, L., Mueller-Dott, S., Trimbour, R., Flores, R. R. O., Argelaguet, R., et al. (2023). Gene regulatory network inference in the era of single-cell multi-omics. *Nat. Rev. Genet.* 24, 739–754. doi: 10.1038/s41576-023-00618-5

- Bailey, T. L., Johnson, J., Grant, C. E., and Noble, W. S. (2015). The MEME suite. *Nucleic Acids Res.* 43, W39–W49. doi: 10.1093/nar/gkv416
- Barah, P., B, N.M., Jayavelu, N. D., Sowdhamini, R., Shameer, K., and Bones, A. M. (2016). Transcriptional regulatory networks in *Arabidopsis thaliana* during single and combined stresses. *Nucleic Acids Res.* 44, 3147–3164. doi: 10.1093/nar/gkv1463
- Bartlett, A., O'Malley, R. C., Huang, S. S. C., Galli, M., Nery, J. R., Gallavotti, A., et al. (2017). Mapping genome-wide transcription-factor binding sites using DAP-seq. *Nat. Protoc.* 12, 1659–1672. doi: 10.1038/nprot.2017.055
- Berger, M. F., and Bulyk, M. L. (2006). Protein binding microarrays (PBMs) for rapid, high-throughput characterization of the sequence specificities of DNA binding proteins. *Methods Mol. Biol. (Clifton N.J.)* 338, 245–260. doi: 10.1385/1-59745-097-9:245
- Berggard, T., Linse, S., and James, P. (2007). Methods for the detection and analysis of protein-protein interactions. *Proteomics* 7, 2833–2842. doi: 10.1002/pmic.200700131
- Borrill, P., Harrington, S. A., Simmonds, J., and Uauy, C. (2019). Identification of transcription factors regulating senescence in wheat through gene regulatory network modelling. *Plant Physiol.* 180, 1740–1755. doi: 10.1104/pp.19.00380
- Boyle, A. P., Davis, S., Shulha, H. P., Meltzer, P., Margulies, E. H., Weng, Z., et al. (2008). High-resolution mapping and characterization of open chromatin across the genome. *Cell* 132, 311–322. doi: 10.1016/j.cell.2007.12.014
- Bravo Gonzalez-Blas, C., De Winter, S., Hulselmans, G., Hecker, N., Matetovici, L., Christiaens, V., et al. (2023). SCENIC+: single-cell multiomic inference of enhancers and gene regulatory networks. *Nat. Methods* 20, 1355–1367. doi: 10.1038/s41592-023-01938-4
- Brkljacic, J., and Grotewold, E. (2017). Combinatorial control of plant gene expression. *Biochim. Biophys. Acta Gene Regul. Mech.* 1860, 31–40. doi: 10.1016/j.bbagrm.2016.07.005
- Brooks, M., Cirrone, J., Pasquino, A., Álvarez, J. M., Swift, J., Mittal, S., et al. (2019). Network Walking charts transcriptional dynamics of nitrogen signaling by integrating validated and predicted genome-wide interactions. *Nat. Commun.* 10, 1569. doi: 10.1038/s41467-019-09522-1
- Brooks, M. D., Juang, C. L., Katari, M. S., Alvarez, J. M., Pasquino, A., Shih, H. J., et al. (2021). ConnecTF: A platform to integrate transcription factor-gene interactions and validate regulatory networks. *Plant Physiol.* 185, 49–66. doi: 10.1093/plphys/kiab012
- Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y., and Greenleaf, W. J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* 10, 1213–1218. doi: 10.1038/nmeth.2688
- Buenrostro, J. D., Wu, B., Litzenburger, U. M., Ruff, D., Gonzales, M. L., Snyder, M. P., et al. (2015). Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* 523, 486–490. doi: 10.1038/nature14590
- Carthew, R. W. (2021). Gene regulation and cellular metabolism: an essential partnership. *Trends Genet.* 37, 389–400. doi: 10.1016/j.tig.2020.09.018
- Cassan, O., Pimpare, L.-L., Dubos, C., Gojon, A., Bach, L., Lebre, S., et al. (2023). A gene regulatory network in *Arabidopsis* roots reveals features and regulators of the plant response to elevated CO<sub>2</sub>. *New Phytol.* 239, 992–1004. doi: 10.1111/nph.18788
- Cazares, T. A., Rizvi, F. W., Iyer, B., Chen, X., Kotliar, M., Bejjani, A. T., et al. (2023). maxATAC: Genome-scale transcription-factor binding prediction from ATAC-seq with deep neural networks. *PLoS Comput. Biol.* 19, e1010863. doi: 10.1371/journal.pcbi.1010863
- Chan, T. E., Stumpf, M. P. H., and Babbie, A. C. (2017). Gene regulatory network inference from single-cell data using multivariate information measures. *Cell Syst.* 5, 251–267.e253. doi: 10.1016/j.cels.2017.08.014
- Chen, K., Chen, J., Pi, X., Huang, L. J., and Li, N. (2023a). Isolation, purification, and application of protoplasts and transient expression systems in plants. *Int. J. Mol. Sci.* 24, 16892. doi: 10.3390/ijms242316892
- Chen, Y., Guo, Y., Guan, P., Wang, Y., Wang, X., Wang, Z., et al. (2023b). A wheat integrative regulatory network from large-scale complementary functional datasets enables trait-associated gene discovery for crop improvement. *Mol. Plant* 16, 393–414. doi: 10.1016/j.molp.2022.12.019
- Chen, G., Ning, B. T., and Shi, T. L. (2019). Single-cell RNA-seq technologies and related computational data analysis. *Front. Genet.* 10. doi: 10.3389/fgenet.2019.00317
- Chen, D., Yan, W., Fu, L. Y., and Kaufmann, K. (2018). Architecture of gene regulatory networks controlling flower development in *Arabidopsis thaliana*. *Nat. Commun.* 9, 4534. doi: 10.1038/s41467-018-06772-3
- Cheng, C. Y., Li, Y., Varala, K., Bubert, J., Huang, J., Kim, G. J., et al. (2021). Evolutionarily informed machine learning enhances the power of predictive gene-to-phenotype relationships. *Nat. Commun.* 12, 5627. doi: 10.1038/s41467-021-25893-w
- Cheng, H., Liu, L., Zhou, Y., Deng, K., Ge, Y., and Hu, X. (2023). TSPTFBS 2.0: trans-species prediction of transcription factor binding sites and identification of their core motifs in plants. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1175837
- Chow, C. N., Lee, T. Y., Hung, Y. C., Li, G. Z., Tseng, K. C., Liu, Y. H., et al. (2019). PlantPAN3.0: a new and updated resource for reconstructing transcriptional regulatory networks from ChIP-seq experiments in plants. *Nucleic Acids Res.* 47, D1155–D1163. doi: 10.1093/nar/gky1081
- Chow, C. N., Yang, C. W., Wu, N. Y., Wang, H. T., Tseng, K. C., Chiu, Y. H., et al. (2024). PlantPAN 4.0: updated database for identifying conserved non-coding sequences and exploring dynamic transcriptional regulation in plant promoters. *Nucleic Acids Res.* 52, D1569–D1578. doi: 10.1093/nar/gkad945
- Chow, C. N., Zheng, H. Q., Wu, N. Y., Chien, C. H., Huang, H. D., Lee, T. Y., et al. (2016). PlantPAN 2.0: an update of plant promoter analysis navigator for reconstructing transcriptional regulatory networks in plants. *Nucleic Acids Res.* 44, D1154–D1160. doi: 10.1093/nar/gkv1035
- Chowdhury, R. H., Eti, F. S., Ahmed, R., Gupta, S. D., Jhan, P. K., Islam, T., et al. (2023). Drought-responsive genes in tomato: meta-analysis of gene expression using machine learning. *Sci. Rep.* 13, 19374. doi: 10.1038/s41598-023-45942-2
- Christie, N., Myburg, A. A., Joubert, F., Murray, S. L., Carstens, M., Lin, Y. C., et al. (2017). Systems genetics reveals a transcriptional network associated with susceptibility in the maize-grey leaf spot pathosystem. *Plant J.* 89, 746–763. doi: 10.1111/tj.13419
- Consortium, A. (2011). Evidence for network evolution in an *Arabidopsis* interactome map. *Science* 333, 601–607. doi: 10.1126/science.1203877
- Cui, S., Youn, E., Lee, J., and Maas, S. J. (2014). An improved systematic approach to predicting transcription factor target genes using support vector machine. *PLoS One* 9, e94519. doi: 10.1371/journal.pone.0094519
- Dai, X., Tu, X., Du, B., Dong, P., Sun, S., Wang, X., et al. (2022). Chromatin and regulatory differentiation between bundle sheath and mesophyll cells in maize. *Plant J.* 109, 675–692. doi: 10.1111/tj.15586
- de Abreu, E. L. F., Li, K., Wen, W., Yan, J., Nikoloski, Z., Willmitzer, L., et al. (2018). Unraveling lipid metabolism in maize with time-resolved multi-omics data. *Plant J.* 93, 1102–1115. doi: 10.1111/tj.13833
- De Bodt, S., Hollunder, J., Nelissen, H., Meulemeester, N., and Inzé, D. (2012). CORNET 2.0: integrating plant coexpression, protein-protein interactions, regulatory interactions, gene associations and functional annotations. *New Phytol.* 195, 707–720. doi: 10.1111/j.1469-8137.2012.04184.x
- De Clercq, I., Van de Velde, J., Luo, X. P., Liu, L., Storme, V., Van Bel, M., et al. (2021). Integrative inference of transcriptional networks in *Arabidopsis* yields novel ROS signalling regulators. *Nat. Plants* 7, 500–504. doi: 10.1038/s41477-021-00894-1
- de la Fuente, A., Bing, N., Hoeschele, I., and Mendes, P. (2004). Discovery of meaningful associations in genomic data using partial correlation coefficients. *Bioinformatics* 20, 3565–3574. doi: 10.1093/bioinformatics/bth445
- de Luis Balaguer, M. A., Fisher, A. P., Clark, N. M., Fernandez-Espinosa, M. G., Möller, B. K., Weijers, D., et al. (2017). Predicting gene regulatory networks by combining spatial and temporal gene expression data in *Arabidopsis* root stem cells. *Proc. Natl. Acad. Sci. U.S.A.* 114, E7632–e7640. doi: 10.1073/pnas.1707566114
- Depuydt, T., De Rybel, B., and Vandepoele, K. (2023). Charting plant gene functions in the multi-omics and single-cell era. *Trends Plant Sci.* 28, 283–296. doi: 10.1016/j.tplants.2022.09.008
- Depuydt, T., and Vandepoele, K. (2021). Multi-omics network-based functional annotation of unknown *Arabidopsis* genes. *Plant J.* 108, 1193–1212. doi: 10.1111/tj.15507
- Dorrity, M. W., Alexandre, C. M., Hamm, M. O., Vigil, A. L., Fields, S., Queitsch, C., et al. (2021). The regulatory landscape of *Arabidopsis thaliana* roots at single-cell resolution. *Nat. Commun.* 12, 3334. doi: 10.1038/s41467-021-23675-y
- Duren, Z., Chen, X., Jiang, R., Wang, Y., and Wong, W. H. (2017). Modeling gene regulation from paired expression and chromatin accessibility data. *Proc. Natl. Acad. Sci. U.S.A.* 114, E4914–E4923. doi: 10.1073/pnas.1704553114
- Duren, Z., Lu, W. S., Arthur, J. G., Shah, P., Xin, J., Meschi, F., et al. (2021). Sc-compReg enables the comparison of gene regulatory networks between conditions using single-cell data. *Nat. Commun.* 12, 4763. doi: 10.1038/s41467-021-25089-2
- Feng, J. W., Han, L., Liu, H., Xie, W. Z., Liu, H., Li, L., et al. (2023). MaizeNetome: A multi-omics network database for functional genomics in maize. *Mol. Plant* 16, 1229–1231. doi: 10.1016/j.molp.2023.08.002
- Feng, C., Song, C., Song, S., Zhang, G., Yin, M., Zhang, Y., et al. (2024). KnockTF 2.0: a comprehensive gene expression profile database with knockdown/knockout of transcription (co-)factors in multiple species. *Nucleic Acids Res.* 52, D183–D193. doi: 10.1093/nar/gkad1016
- Ferrari, C., Manosalva Perez, N., and Vandepoele, K. (2022). MINI-EX: Integrative inference of single-cell gene regulatory networks in plants. *Mol. Plant* 15, 1807–1824. doi: 10.1016/j.molp.2022.10.016
- Ferraz, R. A. C., Lopes, A. L. G., da Silva, J. A. F., Moreira, D. F. V., Ferreira, M. J. N., and de Almeida Coimbra, S. V. (2021). DNA-protein interaction studies: a historical and comparative analysis. *Plant Methods* 17, 82. doi: 10.1186/s13007-021-00780-z
- Franco-Zorrilla, J. M., Lopez-Vidriero, I., Carrasco, J. L., Godoy, M., Vera, P., and Solano, R. (2014). DNA-binding specificities of plant transcription factors and their potential to define target genes. *Proc. Natl. Acad. Sci. United States America* 111, 2367–2372. doi: 10.1073/pnas.1316278111
- Friedman, N. (2004). Inferring cellular networks using probabilistic graphical models. *Science* 303, 799–805. doi: 10.1126/science.1094068
- Fu, J., McKinley, B., James, B., Chrisler, W., Markillie, L. M., Gaffrey, M. J., et al. (2024). Cell-type-specific transcriptomics uncovers spatial regulatory networks in bioenergy sorghum stems. *Plant J.* doi: 10.1111/tj.16690
- Fu, Y., Xiao, W., Tian, L., Guo, L., Ma, G., Ji, C., et al. (2023). Spatial transcriptomics uncover sucrose post-phloem transport during maize kernel development. *Nat. Commun.* 14, 7191. doi: 10.1038/s41467-023-43006-7

- Fu, L. Y., Zhu, T., Zhou, X., Yu, R., He, Z., Zhang, P., et al. (2022). ChIP-Hub provides an integrative platform for exploring plant regulome. *Nat. Commun.* 13, 3413. doi: 10.1038/s41467-022-30770-1
- Galli, M., Khakhar, A., Lu, Z., Chen, Z., Sen, S., Joshi, T., et al. (2018). The DNA binding landscape of the maize AUXIN RESPONSE FACTOR family. *Nat. Commun.* 9, 4526. doi: 10.1038/s41467-018-06977-6
- Gaudinier, A., Rodriguez-Medina, J., Zhang, L. F., Olson, A., Liseron-Monfils, C., Bågman, A. M., et al. (2018). Transcriptional regulation of nitrogen-associated metabolism and growth. *Nature* 563, 259–25+. doi: 10.1038/s41586-018-0656-3
- Gaudinier, A., Zhang, L., Reece-Hoyes, J. S., Taylor-Teeple, M., Pu, L., Liu, Z., et al. (2011). Enhanced Y1H assays for arabidopsis. *Nat. Methods* 8, 1053–1055. doi: 10.1038/nmeth.1750
- Geng, H., Wang, M., Gong, J., Xu, Y., and Ma, S. (2021). An Arabidopsis expression predictor enables inference of transcriptional regulators for gene modules. *Plant J.* 107, 597–612. doi: 10.1111/tpj.15315
- Gomez-Cano, F., Rodriguez, J., Zhou, P., Chu, Y. H., Magnusson, E., Gomez-Cano, L., et al. (2024). Prioritizing metabolic gene regulators through multi-omic network integration in maize. *bioRxiv* 2024, 2024.02.26.582075. doi: 10.1101/2024.02.26.582075
- Grant, C. E., Bailey, T. L., and Noble, W. S. (2011). FIMO: scanning for occurrences of a given motif. *Bioinformatics* 27, 1017–1018. doi: 10.1093/bioinformatics/btr064
- Grotewold, E. (2008). Transcription factors for predictive plant metabolic engineering: are we there yet? *Curr. Opin. Biotechnol.* 19, 138–144. doi: 10.1016/j.copbio.2008.02.002
- Guo, C., Huang, Z., Chen, J., Yu, G., Wang, Y., and Wang, X. (2024). Identification of novel regulators of leaf senescence using a deep learning model. *Plants (Basel)* 13, 1276. doi: 10.3390/plants13091276
- Gupta, O. P., Deshmukh, R., Kumar, A., Singh, S. K., Sharma, P., Ram, S., et al. (2022). From gene to biomolecular networks: a review of evidences for understanding complex biological function in plants. *Curr. Opin. Biotechnol.* 74, 66–74. doi: 10.1016/j.copbio.2021.10.023
- Gupta, S., Gupta, V., Singh, V., and Varadwaj, P. K. (2018). Extrapolation of significant genes and transcriptional regulatory networks involved in *Zea mays* in response in UV-B stress. *Genes Genomics* 40, 973–990. doi: 10.1007/s13258-018-0705-1
- Han, L., Zhong, W., Qian, J., Jin, M., Tian, P., Zhu, W., et al. (2023). A multi-omics integrative network map of maize. *Nat. Genet.* 55, 144–153. doi: 10.1038/s41588-022-01262-1
- Hao, Y., Stuart, T., Kowalski, M. H., Choudhary, S., Hoffman, P., Hartman, A., et al. (2024). Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat. Biotechnol.* 42, 293–304. doi: 10.1038/s41587-023-01767-y
- Harrington, S. A., Backhaus, A. E., Singh, A., Hassani-Pak, K., and Uauy, C. (2020). The wheat GENIE3 network provides biologically-relevant information in polyploid wheat. *G3-Genes Genomes Genet.* 10, 3675–3686. doi: 10.1534/g3.120.401436
- He, Q., Johnston, J., and Zeitlinger, J. (2015). ChIP-nexus enables improved detection of *in vivo* transcription factor binding footprints. *Nat. Biotechnol.* 33, 395–401. doi: 10.1038/nbt.3121
- He, J., Wang, J., and Zhang, Z. (2024). Toward unveiling transcriptome dynamics and regulatory modules at the maternal/filial interface of developing maize kernel. *Plant J.* Online ahead. doi: 10.1111/tpj.16733
- Hecker, M., Lambeck, S., Toepfer, S., van Someren, E., and Guthke, R. (2009). Gene regulatory network inference: data integration in dynamic models—a review. *Biosystems* 96, 86–103. doi: 10.1016/j.biosystems.2008.12.004
- Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., et al. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* 38, 576–589. doi: 10.1016/j.molcel.2010.05.004
- Hickman, R., Van Verk, M. C., Van Dijken, A. J. H., Mendes, M. P., Vroegop-Vos, I. A., Caarls, L., et al. (2017). Architecture and dynamics of the jasmonic acid gene regulatory network. *Plant Cell* 29, 2086–2105. doi: 10.1105/tpc.16.00958
- Hu, J., Zhuang, Y., Li, X., Li, X., Sun, C., Ding, Z., et al. (2022). Time-series transcriptome comparison reveals the gene regulation network under salt stress in soybean (*Glycine max*) roots. *BMC Plant Biol.* 22, 157. doi: 10.1186/s12870-022-03541-9
- Huang, J., Vendramin, S., Shi, L., and McGinnis, K. M. (2017). Construction and optimization of a large gene coexpression network in maize using RNA-seq data. *Plant Physiol.* 175, 568–583. doi: 10.1104/pp.17.00825
- Huang, J., Zheng, J., Yuan, H., and McGinnis, K. (2018). Distinct tissue-specific transcriptional regulation revealed by gene regulatory networks in maize. *BMC Plant Biol.* 18, 111. doi: 10.1186/s12870-018-1329-y
- Hume, M. A., Barrera, L. A., Gisselbrecht, S. S., and Bulyk, M. L. (2015). UniPROBE, update 2015: new tools and content for the online database of protein-binding microarray data on protein-DNA interactions. *Nucleic Acids Res.* 43, D117–D122. doi: 10.1093/nar/gku1045
- Huynh-Thu, V. A., and Geurts, P. (2018). dynGENIE3: dynamical GENIE3 for the inference of gene networks from time series expression data. *Sci. Rep.* 8, 3384. doi: 10.1038/s41598-018-21715-0
- Huynh-Thu, V. A., Irrthum, A., Wehenkel, L., and Geurts, P. (2010). Inferring regulatory networks from expression data using tree-based methods. *PLoS One* 5, 10. doi: 10.1371/journal.pone.0012776
- Huynh-Thu, V. A., and Sanguinetti, G. (2019). “Gene Regulatory Network Inference: An Introductory Survey,” in *Gene Regulatory Networks: Methods and Protocols*, Eds. S. G. and V. A. Huynh-Thu (Springer New York, New York, NY), 1–23. doi: 10.1007/978-1-4939-8882-2\_1
- Jackson, C. A., Castro, D. M., Saldi, G. A., Bonneau, R., and Gresham, D. (2020). Gene regulatory network reconstruction using single-cell RNA sequencing of barcoded genotypes in diverse environments. *Elife* 9, e51254. doi: 10.7554/eLife.51254
- Jansen, C., Ramirez, R. N., El-Ali, N. C., Gomez-Cabrero, D., Tegner, J., Merkschlager, M., et al. (2019). Building gene regulatory networks from scATAC-seq and scRNA-seq using Linked Self Organizing Maps. *PLoS Comput. Biol.* 15, e1006555. doi: 10.1371/journal.pcbi.1006555
- Jasper, S., Nicolás Manosalva, P., Thomas, D., Klaas, V., and Svitlana, L. (2023). MINI-EX version 2: cell-type-specific gene regulatory network inference using an integrative single-cell transcriptomics approach. *bioRxiv* 2023, 2012.2024.573246. doi: 10.1101/2023.12.24.573246
- Jayaram, N., Usvyat, D., and R Martin, A. C. (2016). Evaluating tools for transcription factor binding site prediction. *BMC Bioinf.* 17, 547–547. doi: 10.1186/s12859-016-1298-9
- Ji, C., Xu, L., Li, Y., Fu, Y., Li, S., Wang, Q., et al. (2022). The O2-ZmGRAS11 transcriptional regulatory network orchestrates the coordination of endosperm cell expansion and grain filling in maize. *Mol. Plant* 15, 468–487. doi: 10.1016/j.molp.2021.11.013
- Jiang, L., Gao, Y., Han, L., Zhang, W., and Fan, P. (2023). Designing plant flavonoids: harnessing transcriptional regulation and enzyme variation to enhance yield and diversity. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1220062
- Jiang, Y., Harigaya, Y., Zhang, Z., Zhang, H., Zang, C., and Zhang, N. R. (2022b). Nonparametric single-cell multiomic characterization of trio relationships between transcription factors, target genes, and cis-regulatory regions. *Cell Syst.* 13, 737–751. doi: 10.1016/j.cels.2022.08.004
- Jiang, J., Lyu, P., Li, J., Huang, S., Tao, J., Blackshaw, S., et al. (2022a). iReNA: Integrated regulatory network analysis of single-cell transcriptomes and chromatin accessibility profiles. *iScience* 25, 105359. doi: 10.1016/j.isci.2022.105359
- Jin, J., Tian, F., Yang, D.-C., Meng, Y.-Q., Kong, L., Luo, J., et al. (2017). PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res.* 45, D1040–D1045. doi: 10.1093/nar/gkw982
- Jin, J., Zhang, H., Kong, L., Gao, G., and Luo, J. (2014). PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors. *Nucleic Acids Res.* 42, D1182–D1187. doi: 10.1093/nar/gkt1016
- Johnson, D. S., Mortazavi, A., Myers, R. M., and Wold, B. (2007). Genome-wide mapping of *in vivo* protein-DNA interactions. *Science* 316, 1497–1502. doi: 10.1126/science.1141319
- Joly-Lopez, Z., Platts, A. E., Gulko, B., Choi, J. Y., Groen, S. C., Zhong, X., et al. (2020). An inferred fitness consequence map of the rice genome. *Nat. Plants* 6, 119–130. doi: 10.1038/s41477-019-0589-3
- Kang, Y., Patel, N. R., Shively, C., Recio, P. S., Chen, X., Wrnank, B. J., et al. (2020). Dual threshold optimization and network inference reveal convergent evidence from TF binding locations and TF perturbation responses. *Genome Res.* 30, 459–471. doi: 10.1101/gr.259655.119
- Khamis, A. M., Motwalli, O., Oliva, R., Jankovic, B. R., Medvedeva, Y. A., Ashoor, H., et al. (2018). A novel method for improved accuracy of transcription factor binding site prediction. *Nucleic Acids Res.* 46, e72. doi: 10.1093/nar/gky237
- Kim, J. S., Chae, S., Jun, K. M., Lee, G. S., Jeon, J. S., Kim, K. D., et al. (2021a). Rice protein-binding microarrays: a tool to detect cis-acting elements near promoter regions in rice. *Planta* 253, 40. doi: 10.1007/s00425-021-03572-w
- Kim, E., Hwang, S., and Lee, I. (2017a). SoyNet: a database of co-functional networks for soybean *Glycine max*. *Nucleic Acids Res.* 45, D1082–D1089. doi: 10.1093/nar/gkw704
- Kim, H., Kim, B. S., Shim, J. E., Hwang, S., Yang, S., Kim, E., et al. (2017b). TomatoNet: A genome-wide co-functional network for unveiling complex traits of tomato, a model crop for fleshy fruits. *Mol. Plant* 10, 652–655. doi: 10.1016/j.molp.2016.11.010
- Kim, D., Tran, A., Kim, H. J., Lin, Y., Yang, J. Y. H., and Yang, P. (2023). Gene regulatory network reconstruction: harnessing the power of single-cell multi-omic data. *NPJ Syst. Biol. Appl.* 9, 51. doi: 10.1038/s41540-023-00312-6
- Kim, S. B., Van den Broeck, L., Karre, S., Choi, H., Christensen, S. A., Wang, G. F., et al. (2021b). Analysis of the transcriptomic, metabolic, and gene regulatory responses to Puccinia sorghi in maize. *Mol. Plant Pathol.* 22, 465–479. doi: 10.1111/mpp.13040
- Korhonen, J. H., Palin, K., Taipale, J., and Ukkonen, E. (2017). Fast motif matching revisited: high-order PWMs, SNPs and indels. *Bioinformatics* 33, 514–521. doi: 10.1093/bioinformatics/btw683
- Kuang, J. F., Wu, C. J., Guo, Y. F., Walther, D., Shan, W., Chen, J. Y., et al. (2021). Deciphering transcriptional regulators of banana fruit ripening by regulatory network analysis. *Plant Biotechnol. J.* 19, 477–489. doi: 10.1111/pbi.13477

- Kulkarni, S. R., Vanechoutte, D., Van de Velde, J., and Vandepoele, K. (2018). TF2Network: predicting transcription factor regulators and gene regulatory networks in Arabidopsis using publicly available binding site information. *Nucleic Acids Res.* 46, e31. doi: 10.1093/nar/gkx1279
- Lai, X. L., Stigliani, A., Lucas, J., Hougouvieux, V., Parcy, F., and Zubieta, C. (2020). Genome-wide binding of SEPALLATA3 and AGAMOUS complexes determined by sequential DNA-affinity purification sequencing. *Nucleic Acids Res.* 48, 9637–9648. doi: 10.1093/nar/gkaa729
- Lai, X., Stigliani, A., Vachon, G., Carles, C., Smaczniak, C., Zubieta, C., et al. (2019). Building transcription factor binding site models to understand gene regulation in plants. *Mol. Plant* 12, 743–763. doi: 10.1016/j.molp.2018.10.010
- Lai, X., Wolkenhauer, O., and Vera, J. (2016). Understanding microRNA-mediated gene regulatory networks through mathematical modelling. *Nucleic Acids Res.* 44, 6019–6035. doi: 10.1093/nar/gkw550
- Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf.* 9, 559. doi: 10.1186/1471-2105-9-559
- Lee, T., Hwang, S., Kim, C. Y., Shim, H., Kim, H., Ronald, P. C., et al. (2017b). WheatNet: a genome-scale functional network for hexaploid bread wheat, triticum aestivum. *Mol. Plant* 10, 1133–1136. doi: 10.1016/j.molp.2017.04.006
- Lee, J. H., Jin, S., Kim, S. Y., Kim, W., and Ahn, J. H. (2017a). A fast, efficient chromatin immunoprecipitation method for studying protein-DNA binding in Arabidopsis mesophyll protoplasts. *Plant Methods* 13, 42. doi: 10.1186/s13007-017-0192-4
- Lee, T., Lee, S., Yang, S., and Lee, I. (2019). MaizeNet: a co-functional network for network-assisted systems genetics in Zea mays. *Plant J.* 99, 571–582. doi: 10.1111/tj.14341
- Lee, T., Oh, T., Yang, S., Shin, J., Hwang, S., Kim, C. Y., et al. (2015a). RiceNet v2: an improved network prioritization server for rice genes. *Nucleic Acids Res.* 43, W122–W127. doi: 10.1093/nar/gkv253
- Lee, W., Park, B., and Han, K. (2017c). Sequence-based prediction of putative transcription factor binding sites in DNA sequences of any length. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 15, 1461–1469. doi: 10.1109/TCBB.2017.2773075
- Lee, L., Seo, Y. S., Coltrane, D., Hwang, S., Oh, T., Marcotte, E. M., et al. (2011). Genetic dissection of the biotic stress response using a genome-scale gene network for rice. *Proc. Natl. Acad. Sci. U.S.A.* 108, 18548–18553. doi: 10.1073/pnas.1110384108
- Lee, T., Yang, S., Kim, E., Ko, Y., Hwang, S., Shin, J., et al. (2015b). AraNet v2: an improved database of co-functional gene networks for the study of Arabidopsis thaliana and 27 other nonmodel plant species. *Nucleic Acids Res.* 43, D996–1002. doi: 10.1093/nar/gku1053
- Li, Z. Q., Hu, Y. H., Ma, X. L., Da, L., She, J. J., Liu, Y., et al. (2023b). WheatCENet: A database for comparative co-expression network analysis of allohexaploid wheat and its progenitors. *Genomics Proteomics Bioinf.* 21, 324–336. doi: 10.1016/j.gpb.2022.04.007
- Li, Y., Pearl, S. A., and Jackson, S. A. (2015). Gene networks in plant biology: approaches in reconstruction and analysis. *Trends Plant Sci.* 20, 664–675. doi: 10.1016/j.tplants.2015.06.013
- Li, M., Yao, T., Lin, W., Hincley, W. E., Galli, M., Muchero, W., et al. (2023a). Double DAP-seq uncovered synergistic DNA binding of interacting bZIP transcription factors. *Nat. Commun.* 14, 2600. doi: 10.1038/s41467-023-38096-2
- Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 289–293. doi: 10.1126/science.1181369
- Lin, X., Xu, Y., Wang, D., Yang, Y., Zhang, X., Bie, X., et al. (2024). Systematic identification of wheat spike developmental regulators by integrated multi-omics, transcriptional network, GWAS, and genetic analyses. *Mol. Plant* 17, 438–459. doi: 10.1016/j.molp.2024.01.010
- Lin, C. T., Xu, T., Xing, S. L., Zhao, L., Sun, R. Z., Liu, Y., et al. (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 Cell Physiol.* 60, 2707–2719. doi: 10.1093/pcp/pcz160
- Lin, H., Yu, J., Pearce, S. P., Zhang, D., and Wilson, Z. A. (2017). RiceAntherNet: a gene co-expression network for identifying anther and pollen development genes. *Plant J.* 92, 1076–1091. doi: 10.1111/tj.13744
- Liu, X., Bie, X. M., Lin, X., Li, M., Wang, H., Zhang, X., et al. (2023b). Uncovering the transcriptional regulatory network involved in boosting wheat regeneration and transformation. *Nat. Plants* 9, 908–925. doi: 10.1038/s41477-023-01406-z
- Liu, C., Huang, H., and Yang, P. (2023a). Multi-task learning from multimodal single-cell omics with Matilda. *Nucleic Acids Res.* 51, e45. doi: 10.1093/nar/gkad157
- Liu, S., Liu, Y., Zhao, J., Cai, S., Qian, H., Zuo, K., et al. (2017). A computational interactome for prioritizing genes associated with complex agronomic traits in rice (Oryza sativa). *Plant J.* 90, 177–188. doi: 10.1111/tj.13475
- Liu, Y., Xie, S., and Yu, J. (2016). Genome-wide analysis of the lysine biosynthesis pathway network during maize seed development. *PLoS One* 11, e0148287. doi: 10.1371/journal.pone.0148287
- Lou, S., Li, T., Kong, X., Zhang, J., Liu, J., Lee, D., et al. (2020). TopicNet: a framework for measuring transcriptional regulatory network change. *Bioinformatics* 36, i474–i481. doi: 10.1093/bioinformatics/btaa403
- Loudya, N., Mishra, P., Takahagi, K., Uehara-Yamaguchi, Y., Inoue, K., Bogre, L., et al. (2021). Cellular and transcriptomic analyses reveal two-staged chloroplast biogenesis underpinning photosynthesis build-up in the wheat leaf. *Genome Biol.* 22, 30. doi: 10.1186/s13059-021-02366-3
- Ma, S., Ding, Z., and Li, P. (2017). Maize network analysis revealed gene modules involved in development, nutrients utilization, metabolism, and stress response. *BMC Plant Biol.* 17, 131. doi: 10.1186/s12870-017-1077-4
- Ma, X., Feng, L., Tao, A., Zenda, T., He, Y., Zhang, D., et al. (2023b). Identification and validation of seed dormancy loci and candidate genes and construction of regulatory networks by WGCNA in maize introgression lines. *Theor. Appl. Genet.* 136, 259. doi: 10.1007/s00122-023-04495-8
- Ma, J., Jia, B., Bian, Y., Pei, W., Song, J., Wu, M., et al. (2024). Genomic and co-expression network analyses reveal candidate genes for oil accumulation based on an introgression population in Upland cotton (Gossypium hirsutum). *Theor. Appl. Genet.* 137, 23. doi: 10.1007/s00122-023-04527-3
- Ma, A., Wang, X., Li, J., Wang, C., Xiao, T., Liu, Y., et al. (2023a). Single-cell biological network inference using a heterogeneous graph transformer. *Nat. Commun.* 14, 964. doi: 10.1038/s41467-023-36559-0
- Ma, S. W., Wang, M., Wu, J. H., Guo, W. L., Chen, Y. M., Li, G. W., et al. (2021). WheatOmics: A platform combining multiple omics data to accelerate functional genomics studies in wheat. *Mol. Plant* 14, 1965–1968. doi: 10.1016/j.molp.2021.10.006
- Ma, Q., and Xu, D. (2022). Deep learning shapes single-cell data analysis. *Nat. Rev. Mol. Cell Biol.* 23, 303–304. doi: 10.1038/s41580-022-00466-x
- Mahood, E. H., Kruse, L. H., and Moghe, G. D. (2020). Machine learning: A powerful tool for gene function prediction in plants. *Appl. Plant Sci.* 8, e11376. doi: 10.1002/aps3.11376
- Manosalva Perez, N., Ferrari, C., Engelhorn, J., Depuydt, T., Nelissen, H., Hartwig, T., et al. (2024). MINI-AC: inference of plant gene regulatory networks using bulk or single-cell accessible chromatin profiles. *Plant J.* 117, 280–301. doi: 10.1111/tj.16483
- Marand, A. P., Chen, Z., Gallavotti, A., and Schmitz, R. J. (2021). A cis-regulatory atlas in maize at single-cell resolution. *Cell* 184, 3041–3055 e3021. doi: 10.1016/j.cell.2021.04.014
- Marbach, D., Costello, J. C., Kuffner, R., Vega, N. M., Prill, R. J., Camacho, D. M., et al. (2012a). Wisdom of crowds for robust gene network inference. *Nat. Methods* 9, 796–804. doi: 10.1038/nmeth.2016
- Marbach, D., Roy, S., Ay, F., Meyer, P. E., Candeias, R., Kahveci, T., et al. (2012b). Predictive regulatory models in Drosophila melanogaster by integrative inference of transcriptional networks. *Genome Res.* 22, 1334–1349. doi: 10.1101/gr.127191.111
- Matsumoto, H., Kiryu, H., Furusawa, C., Ko, M. S. H., Ko, S. B. H., Gouda, N., et al. (2017). SCODE: an efficient regulatory network inference algorithm from single-cell RNA-Seq during differentiation. *Bioinformatics* 33, 2314–2321. doi: 10.1093/bioinformatics/btx194
- Matys, V., Fricke, E., Geffers, R., Gossling, E., Haubrock, M., Hehl, R., et al. (2003). TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res.* 31, 374–378. doi: 10.1093/nar/gkg108
- McWhite, C. D., Papoulas, O., Drew, K., Cox, R. M., June, V., Dong, O. X., et al. (2020). A pan-plant protein complex map reveals deep conservation and novel assemblies. *Cell* 181, 460–474 e414. doi: 10.1016/j.cell.2020.02.049
- Meers, M. P., Bryson, T. D., Henikoff, J. G., and Henikoff, S. (2019). Improved CUT&RUN chromatin profiling tools. *Elife* 8, e46314. doi: 10.7554/eLife.46314.018
- Mieczkowski, J., Cook, A., Bowman, S. K., Mueller, B., Alver, B. H., Kundu, S., et al. (2016). MNase titration reveals differences between nucleosome occupancy and chromatin accessibility. *Nat. Commun.* 7, 11485. doi: 10.1038/ncomms11485
- Min, S., Lee, B., and Yoon, S. (2017). Deep learning in bioinformatics. *Brief Bioinform.* 18, 851–869. doi: 10.1093/bib/bbw068
- Moerman, T., Aibar Santos, S., Bravo González-Blas, C., Simm, J., Moreau, Y., Aerts, J., et al. (2019). GRNBoost2 and Arboreto: efficient and scalable inference of gene regulatory networks. *Bioinformatics* 35, 2159–2161. doi: 10.1093/bioinformatics/bty916
- Morin, A., Chu, E. C.-P., Sharma, A., Adrian-Hamazaki, A., and Pavlidis, P. (2023). Characterizing the targets of transcription regulators by aggregating ChIP-seq and perturbation expression data sets. *Genome Res.* 33, 763–778. doi: 10.1101/gr.277273.122
- Movahedi, S., Van de Peer, Y., and Vandepoele, K. (2011). Comparative network analysis reveals that tissue specificity and gene function are important factors influencing the mode of expression evolution in Arabidopsis and rice. *Plant Physiol.* 156, 1316–1330. doi: 10.1104/pp.111.177865
- Nietzsche, M., Landgraf, R., Tohge, T., and Bornke, F. (2016). A protein-protein interaction network linking the energy-sensor kinase SnRK1 to multiple signaling pathways in Arabidopsis thaliana. *Curr. Plant Biol.* 5, 36–44. doi: 10.1016/j.cpb.2015.10.004
- Nolan, T. M., Vukasinovic, N., Hsu, C. W., Zhang, J., Vanhoutte, I., Shahan, R., et al. (2023). Brassinosteroid gene regulatory networks at cellular resolution in the Arabidopsis root. *Science* 379, eadf4721. doi: 10.1126/science.adf4721
- O'Malley, R. C., Huang, S. S. C., Song, L., Lewsey, M. G., Bartlett, A., Nery, J. R., et al. (2016). Cistrome and epistrome features shape the regulatory DNA landscape. *Cell* 165, 1280–1292. doi: 10.1016/j.cell.2016.04.038

- Obayashi, T., Aoki, Y., Tadaka, S., Kagaya, Y., and Kinoshita, K. (2018). ATTED-II in 2018: A plant coexpression database based on investigation of the statistical property of the mutual rank index. *Plant Cell Physiol.* 59, 440. doi: 10.1093/pcp/pcx209
- Obayashi, T., Hibara, H., Kagaya, Y., Aoki, Y., and Kinoshita, K. (2022). ATTED-II v1.1: A plant gene coexpression database using a sample balancing technique by subsampling of principal components. *Plant Cell Physiol.* 63, 869–881. doi: 10.1093/pcp/pcac041
- Oughtred, R., Rust, J., Chang, C., Breitkreutz, B. J., Stark, C., Willems, A., et al. (2021). The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Sci.* 30, 187–200. doi: 10.1002/pro.3978
- Ouyang, W., Xiong, D., Li, G., and Li, X. (2020). Unraveling the 3D genome architecture in plants: present and future. *Mol. Plant* 13, 1676–1693. doi: 10.1016/j.molp.2020.10.002
- Parvathaneni, R. K., Bertolini, E., Shamimuzzaman, M., Vera, D. L., Lung, P. Y., Rice, B. R., et al. (2020). The regulatory landscape of early maize inflorescence development. *Genome Biol.* 21, 33. doi: 10.1186/s13059-020-02070-8
- Peng, Y., Xiong, D., Zhao, L., Ouyang, W., Wang, S., Sun, J., et al. (2019). Chromatin interaction maps reveal genetic regulation for quantitative traits in maize. *Nat. Commun.* 10, 2632. doi: 10.1038/s41467-019-10602-5
- Peng, Y., Zuo, W., Zhou, H., Miao, F., Zhang, Y., Qin, Y., et al. (2022). EXPLICIT-Kinase: A gene expression predictor for dissecting the functions of the Arabidopsis kinome. *J. Integr. Plant Biol.* 64, 1374–1393. doi: 10.1111/jipb.13267
- Pfeifer, M., Kugler, K. G., Sandve, S. R., Zhan, B., Rudi, H., Hvidsten, T. R., et al. (2014). Genome interplay in the grain transcriptome of hexaploid bread wheat. *Science* 345, 1250091. doi: 10.1126/science.1250091
- Pranzatelli, T. J. F., Michael, D. G., and Chiorini, J. A. (2018). ATAC2GRN: optimized ATAC-seq and DNase1-seq pipelines for rapid and accurate genome regulatory network inference. *BMC Genomics* 19, 563. doi: 10.1186/s12864-018-4943-z
- Pratapa, A., Jalihal, A. P., Law, J. N., Bharadwaj, A., and Murali, T. M. (2020). Benchmarking algorithms for gene regulatory network inference from single-cell transcriptomic data. *Nat. Methods* 17, 147–154. doi: 10.1038/s41592-019-0690-6
- Puig, R. R., Boddie, P., Khan, A., Castro-Mondragon, J. A., and Mathelier, A. (2021). UniBind: maps of high-confidence direct TF-DNA interactions across nine species. *BMC Genomics* 22, 482. doi: 10.1186/s12864-021-07760-6
- Qin, Q., Fan, J., Zheng, R., Wan, C., Mei, S., Wu, Q., et al. (2020). Lisa: inferring transcriptional regulators through integrative modeling of public chromatin accessibility and ChIP-seq data. *Genome Biol.* 21, 32. doi: 10.1186/s13059-020-1934-6
- Ramirez-González, R. H., Borrill, P., Lang, D., Harrington, S. A., Brinton, J., Venturini, L., et al. (2018). The transcriptional landscape of polyploid wheat. *Science* 361, 662–664. doi: 10.1126/science.aar6089
- Rao, X. L., and Dixon, R. A. (2019). Co-expression networks for plant biology: why and how. *Acta Biochim. Et Biophys. Sin.* 51, 981–988. doi: 10.1093/abbs/gmz080
- Raulusevičiute, I., Riudavets-Puig, R., Blanc-Mathieu, R., Castro-Mondragon, J. A., Ferenc, K., Kumar, V., et al. (2024). JASPAR 2024: 20th anniversary of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* 52, D174–D182. doi: 10.1093/nar/gkad1059
- Reynoso, M. A., Borowsky, A. T., Pauluzzi, G. C., Yeung, E., Zhang, J., Formentin, E., et al. (2022). Gene regulatory networks shape developmental plasticity of root cell types under water extremes in rice. *Dev. Cell* 57, 1177–1192 e1176. doi: 10.1016/j.devcel.2022.04.013
- Rhee, S. Y., Birnbaum, K. D., and Ehrhardt, D. W. (2019). Towards building a plant cell atlas. *Trends Plant Sci.* 24, 303–310. doi: 10.1016/j.tplants.2019.01.006
- Ricci, W. A., Lu, Z. F., Ji, L. X., Marand, A. P., Ethridge, C. L., Murphy, N. G., et al. (2019). Widespread long-range regulatory elements in the maize genome. *Nat. Plants* 5, 1237–1249. doi: 10.1038/s41477-019-0547-0
- Rossi, M. J., Lai, W. K. M., and Pugh, B. F. (2018). Simplified ChIP-exo assays. *Nat. Commun.* 9, 2842. doi: 10.1038/s41467-018-05265-7
- Ruengrichaiya, B., Nukoolkit, C., Kalapanulak, S., and Saitong, T. (2022). Plant-DTI: Extending the landscape of TF protein and DNA interaction in plants by a machine learning-based approach. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.970018
- Santuari, L., Sanchez-Perez, G. F., Luijten, M., Rutjens, B., Terpstra, I., Berke, L., et al. (2016). The PLETHORA gene regulatory network guides growth and cell differentiation in Arabidopsis roots. *Plant Cell* 28, 2937–2951. doi: 10.1105/tpc.16.00656
- Sari, E., Cabral, A. L., Polley, B., Tan, Y. F., Hsueh, E., Konkin, D. J., et al. (2019). Weighted gene co-expression network analysis unveils gene networks associated with the Fusarium head blight resistance in tetraploid wheat. *BMC Genomics* 20, 925. doi: 10.1186/s12864-019-6161-8
- Schmitz, R. J., Grotewold, E., and Stam, M. (2022). Cis-regulatory sequences in plants: Their importance, discovery, and future challenges. *Plant Cell* 34, 718–741. doi: 10.1093/plcell/koab281
- Scofield, S., Murison, A., Jones, A., Fozard, J., Aida, M., Band, L. R., et al. (2018). Coordination of meristem and boundary functions by transcription factors in the SHOOT MERISTEMLESS regulatory network. *Development* 145, dev157081. doi: 10.1242/dev.157081
- Segal, E., Shapira, M., Regev, A., Pe'er, D., Botstein, D., Koller, D., et al. (2003). Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. *Nat. Genet.* 34, 166–176. doi: 10.1038/ng1165
- Sevilla, M., Gaskins, J., and Datta, S. (2020). A sparse Bayesian factor model for the construction of gene co-expression networks from single-cell RNA sequencing count data. *BMC Bioinf.* 21, 361. doi: 10.1186/s12859-020-03707-y
- Serebreni, L., and Stark, A. (2021). Insights into gene regulation: From regulatory genomic elements to DNA-protein and protein-protein interactions. *Curr. Opin. Cell Biol.* 70, 58–66. doi: 10.1016/j.ceb.2020.11.009
- Shi, J., Zhao, B., Zheng, S., Zhang, X., Wang, X., Dong, W., et al. (2021). A phosphate starvation response-centered network regulates mycorrhizal symbiosis. *Cell* 184, 5527–5540 e5518. doi: 10.1016/j.cell.2021.09.030
- Simon, J. M., Giresi, P. G., Davis, I. J., and Lieb, J. D. (2012). Using formaldehyde-assisted isolation of regulatory elements (FAIRE) to isolate active regulatory DNA. *Nat. Protoc.* 7, 256–267. doi: 10.1038/nprot.2011.444
- Sircar, S., Musaddi, M., and Parekh, N. (2022). NetREx: Network-based Rice Expression Analysis Server for abiotic stress conditions. *Database (Oxford)* 2022, baac060. doi: 10.1093/database/baac060
- Slawek, J., and Arodz, T. (2013). ENNET: inferring large gene regulatory networks from expression data using gradient boosting. *BMC Syst. Biol.* 7, 106. doi: 10.1186/1752-0509-7-106
- Smaczniak, C., Angenent, G. C., and Kaufmann, K. (2017). SELEX-seq: A method to determine DNA binding specificities of plant transcription factors. *Methods Mol. Biol.* 1629, 67–82. doi: 10.1007/978-1-4939-7125-1\_6
- Sonawane, A. R., DeMeo, D. L., Quackenbush, J., and Glass, K. (2021). Constructing gene regulatory networks using epigenetic data. *NPJ Syst. Biol. Appl.* 7, 45. doi: 10.1038/s41540-021-00208-3
- Song, L., Langfelder, P., and Horvath, S. (2012). Comparison of co-expression measures: mutual information, correlation, and model based indices. *BMC Bioinf.* 13, 328. doi: 10.1186/1471-2105-13-328
- Song, Q., Lee, J., Akter, S., Grene, R., and Li, S. (2019). Prediction of condition-specific regulatory maps in Arabidopsis using integrated genomic data. *bioRxiv* 565119. doi: 10.1101/565119
- Song, Q., and Li, S. (2023). Identification of plant co-regulatory modules using coReg. *Methods Mol. Biol.* 2594, 217–223. doi: 10.1007/978-1-0716-2815-7
- Song, X. G., Meng, X. B., Guo, H. Y., Cheng, Q., Jing, Y. H., Chen, M. J., et al. (2022). Targeting a gene regulatory element enhances rice grain yield by decoupling panicle number and size. *Nat. Biotechnol.* 40, 1403–1404. doi: 10.1038/s41587-022-01281-7
- Song, Q., Ruffalo, M., and Bar-Joseph, Z. (2023). Using single cell atlas data to reconstruct regulatory networks. *Nucleic Acids Res.* 51, e38. doi: 10.1093/nar/gkad053
- Strader, L., Weijers, D., and Wagner, D. (2022). Plant transcription factors - being in the right place with the right company. *Curr. Opin. Plant Biol.* 65, 102136. doi: 10.1016/j.jpb.2021.102136
- Stuart, J. M., Segal, E., Koller, D., and Kim, S. K. (2003). A gene-coexpression network for global discovery of conserved genetic modules. *Science* 302, 249–255. doi: 10.1126/science.1087447
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* 102, 15545–15550. doi: 10.1073/pnas.0506580102
- Sullivan, A. M., Arsovski, A. A., Lempe, J., Bubb, K. L., Weirauch, M. T., Sabo, P. J., et al. (2014). Mapping and dynamics of regulatory DNA and transcription factor networks in A-thaliana. *Cell Rep.* 8, 2015–2030. doi: 10.1016/j.celrep.2014.08.019
- Tai, Y. L., Liu, C., Yu, S. W., Yang, H., Sun, J. M., Guo, C. X., et al. (2018). Gene co-expression network analysis reveals coordinated regulation of three characteristic secondary biosynthetic pathways in tea plant (*Camellia sinensis*). *BMC Genomics* 19, 616. doi: 10.1186/s12864-018-4999-9
- Tang, Q., Chen, Y., Meyer, C., Geistlinger, T., Lupien, M., Wang, Q., et al. (2011). A comprehensive view of nuclear receptor cancer cistromes. *Cancer Res.* 71, 6940–6947. doi: 10.1158/0008-5472.CAN-11-2091
- Tang, T., Tian, S., Wang, H., Lv, X., Xie, Y., Liu, J., et al. (2023). Wheat-RegNet: An encyclopedia of common wheat hierarchical regulatory networks. *Mol. Plant* 16, 318–321. doi: 10.1016/j.molp.2022.12.018
- Tao, X. Y., Feng, S. L., Zhao, T., and Guan, X. Y. (2020). Efficient chromatin profiling of H3K4me3 modification in cotton using CUT&Tag. *Plant Methods* 16, 120. doi: 10.1186/s13007-020-00664-8
- Tao, S., Liu, P., Shi, Y., Feng, Y., Gao, J., Chen, L., et al. (2022). Single-cell transcriptome and network analyses unveil key transcription factors regulating mesophyll cell development in maize. *Genes (Basel)* 13. doi: 10.3390/genes13020374
- Tao, S., and Zhang, W. (2022). Network and epigenetic characterization of subsets of genes specifically expressed in maize bundle sheath cells. *Comput. Struct. Biotechnol. J.* 20, 3581–3590. doi: 10.1016/j.csbj.2022.07.004
- Thompson, D., Regev, A., and Roy, S. (2015). Comparative analysis of gene regulatory networks: from network reconstruction to evolution. *Annu. Rev. Cell Dev. Biol.* 31, 399–428. doi: 10.1146/annurev-cellbio-100913-012908
- Tian, F., Yang, D. C., Meng, Y. Q., Jin, J. P., and Gao, G. (2020). PlantRegMap: charting functional regulatory maps in plants. *Nucleic Acids Res.* 48, D1104–D1113. doi: 10.1093/nar/gkz1020

- Tu, X. Y., Mejía-Guerra, M. K., Franco, J. A. V., Tzeng, D., Chu, P. Y., Shen, W., et al. (2020). Reconstructing the maize leaf regulatory network using ChIP-seq data of 104 transcription factors. *Nat. Commun.* 11, 5089. doi: 10.1038/s41467-020-18832-8
- Van den Broeck, L., Gordon, M., Inzé, D., Williams, C., and Sozzani, R. (2020). Gene regulatory network inference: connecting plant biology and mathematical modeling. *Front. Genet.* 11. doi: 10.3389/fgene.2020.00457
- Vandepoele, K., Quimbaya, M., Casneuf, T., De Veylder, L., and Van de Peer, Y. (2009). Unraveling transcriptional control in arabidopsis using cis-regulatory elements and coexpression networks. *Plant Physiol.* 150, 535–546. doi: 10.1104/pp.109.136028
- van der Sande, M., Frolich, S., and van Heeringen, S. J. (2023). Computational approaches to understand transcription regulation in development. *Biochem. Soc. Trans.* 51, 1–12. doi: 10.1042/bst20210145
- Van de Sande, B., Flerin, C., Davie, K., De Waegeneer, M., Hulselmans, G., Aibar, S., et al. (2020). A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nat. Protoc.* 15, 2247–2276. doi: 10.1038/s41596-020-0336-2
- van Dijk, D., Sharma, R., Nainys, J., Yim, K., Kathail, P., Carr, A. J., et al. (2018). Recovering gene interactions from single-cell data using data diffusion. *Cell* 174, 716–729.e727. doi: 10.1016/j.cell.2018.05.061
- Varala, K., Marshall-Colón, A., Cirrone, J., Brooks, M. D., Pasquino, A. V., Léran, S., et al. (2018). Temporal transcriptional logic of dynamic regulatory networks underlying nitrogen signaling and use in plants. *Proc. Natl. Acad. Sci. U.S.A.* 115, 6494–6499. doi: 10.1073/pnas.1721487115
- Vignes, M., Vandel, J., Allouche, D., Ramadan-Alban, N., Cierco-Ayrolles, C., Schiex, T., et al. (2011). Gene regulatory network reconstruction using Bayesian networks, the Dantzig Selector, the Lasso and their meta-analysis. *PLoS One* 6, e29165. doi: 10.1371/journal.pone.0029165
- von Mering, C., Krause, R., Snel, B., Cornell, M., Oliver, S. G., Fields, S., et al. (2002). Comparative assessment of large-scale data sets of protein-protein interactions. *Nature* 417, 399–403. doi: 10.1038/nature750
- Walley, J. W., Sartor, R. C., Shen, Z., Schmitz, R. J., Wu, K. J., Urich, M. A., et al. (2016). Integration of omic networks in a developmental atlas of maize. *Science* 353, 814–818. doi: 10.1126/science.aag1125
- Wang, R., Cheng, Y., Ke, X., Zhang, X., Zhang, H., and Huang, J. (2020b). Comparative analysis of salt responsive gene regulatory networks in rice and Arabidopsis. *Comput. Biol. Chem.* 85, 107188. doi: 10.1016/j.compbiolchem.2019.107188
- Wang, X., Li, J., Han, L., Liang, C., Li, J., Shang, X., et al. (2023b). QTG-Miner aids rapid dissection of the genetic base of tassel branch number in maize. *Nat. Commun.* 14, 5232. doi: 10.1038/s41467-023-41022-1
- Wang, G., Li, X., Shen, W., Li, M. W., Huang, M., Zhang, J., et al. (2022a). The chromatin accessibility landscape of pistils and anthers in rice. *Plant Physiol.* 190, 2797–2811. doi: 10.1093/plphys/kiac448
- Wang, S., Sun, H., Ma, J., Zang, C., Wang, C., Wang, J., et al. (2013). Target analysis by integration of transcriptome and ChIP-seq data with BETA. *Nat. Protoc.* 8, 2502–2515. doi: 10.1038/nprot.2013.150
- Wang, L., Trasanidis, N., Wu, T., Dong, G., Hu, M., Bauer, D. E., et al. (2023a). Dictys: dynamic gene regulatory network dissects developmental continuum with single-cell multiomics. *Nat. Methods* 20, 1368–1378. doi: 10.1038/s41592-023-01971-3
- Wang, J., Wang, Y., Lü, T., Yang, X., Liu, J., Dong, Y., et al. (2022b). An efficient and universal protoplast isolation protocol suitable for transient gene expression analysis and single-cell RNA sequencing. *Int. J. Mol. Sci.* 23, 3419. doi: 10.3390/ijms23073419
- Wang, H., Wu, G., Zhao, B., Wang, B., Lang, Z., Zhang, C., et al. (2016). Regulatory modules controlling early shade avoidance response in maize seedlings. *BMC Genomics* 17, 269. doi: 10.1186/s12864-016-2593-6
- Wang, Y. G., Yu, H. P., Tian, C. H., Sajjad, M., Gao, C. C., Tong, Y. P., et al. (2017). Transcriptome association identifies regulators of wheat spike architecture. *Plant Physiol.* 175, 746–757. doi: 10.1104/pp.17.00694
- Wang, Q., Zeng, X. N., Song, Q. L., Sun, Y., Feng, Y. J., and Lai, Y. C. (2020a). Identification of key genes and modules in response to Cadmium stress in different rice varieties and stem nodes by weighted gene co-expression network analysis. *Sci. Rep.* 10, 9525. doi: 10.1038/s41598-020-66132-4
- Weirauch, M. T., Yang, A., Albu, M., Cote, A. G., Montenegro-Montero, A., Drewe, P., et al. (2014). Determination and inference of eukaryotic transcription factor sequence specificity. *Cell* 158, 1431–1443. doi: 10.1016/j.cell.2014.08.009
- Weston, D. J., Karve, A. A., Gunter, L. E., Jawdy, S. S., Yang, X., Allen, S. M., et al. (2011). Comparative physiology and transcriptional networks underlying the heat shock response in *Populus trichocarpa*, *Arabidopsis thaliana* and *Glycine max*. *Plant Cell Environ.* 34, 1488–1506. doi: 10.1111/j.1365-3040.2011.02347.x
- Wisecaver, J. H., Borowsky, A. T., Tzin, V., Jander, G., Kliebenstein, D. J., and Rokas, A. (2017). A global coexpression network approach for connecting genes to specialized metabolic pathways in plants. *Plant Cell* 29, 944–959. doi: 10.1105/tpc.17.00009
- Wolfe, C. J., Kohane, I. S., and Butte, A. J. (2005). Systematic survey reveals general applicability of "guilt-by-association" within gene coexpression networks. *BMC Bioinf.* 6, 227. doi: 10.1186/1471-2105-6-227
- Worsley Hunt, R., and Wasserman, W. W. (2014). Non-targeted transcription factors motifs are a systemic component of ChIP-seq datasets. *Genome Biol.* 15, 412. doi: 10.1186/s13059-014-0412-4
- Wu, L., Luo, Z., Shi, Y., Jiang, Y., Li, R., Miao, X., et al. (2022). A cost-effective tsCUT&Tag method for profiling transcription factor binding landscape. *J. Integr. Plant Biol.* 64, 2033–2038. doi: 10.1111/jipb.13354
- Xiang, D. Q., Quilichini, T. D., Liu, Z. Y., Gao, P., Pan, Y. L., Li, Q., et al. (2019). The transcriptional landscape of polyploid wheats and their diploid ancestors during embryogenesis and grain development. *Plant Cell* 31, 2888–2911. doi: 10.1105/tpc.19.00397
- Xie, Y., Jiang, S., Li, L., Yu, X., Wang, Y., Luo, C., et al. (2020). Single-cell RNA sequencing efficiently predicts transcription factor targets in plants. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.603302
- Xie, L., Liu, M., Zhao, L., Cao, K., Wang, P., Xu, W., et al. (2021). RiceENCODE: A comprehensive epigenomic database as a rice Encyclopedia of DNA Elements. *Mol. Plant* 14, 1604–1606. doi: 10.1016/j.molp.2021.08.018
- Xing, S. P., Wallmeroth, N., Berendzen, K. W., and Grefen, C. (2016). Techniques for the analysis of protein-protein interactions in vivo. *Plant Physiol.* 171, 727–758. doi: 10.1104/pp.16.00470
- Xiong, W., Wang, C., Zhang, X., Yang, Q., Shao, R., Lai, J., et al. (2017). Highly interwoven communities of a gene regulatory network unveil topologically important genes for maize seed development. *Plant J.* 92, 1143–1156. doi: 10.1111/tpj.13750
- Xu, X., Crow, M., Rice, B. R., Li, F., Harris, B., Liu, L., et al. (2021). Single-cell RNA sequencing of developing maize ears facilitates functional analysis and trait candidate gene discovery. *Dev. Cell* 56, 557–568.e556. doi: 10.1016/j.devcel.2020.12.015
- Xu, W., Yang, W., Zhang, Y., Chen, Y., Hong, N., Zhang, Q., et al. (2022b). ISSAAC-seq enables sensitive and flexible multimodal profiling of chromatin accessibility and gene expression in single cells. *Nat. Methods* 19, 1243–1249. doi: 10.1038/s41592-022-01601-4
- Xu, J., Zhu, J., Lin, Y., Zhu, H., Tang, L., Wang, X., et al. (2022a). Comparative transcriptome and weighted correlation network analyses reveal candidate genes involved in chlorogenic acid biosynthesis in sweet potato. *Sci. Rep.* 12, 2770. doi: 10.1038/s41598-022-06794-4
- Yang, Y., Kong, Q., Lim, A. R. Q., Lu, S., Zhao, H., Guo, L., et al. (2022). Transcriptional regulation of oil biosynthesis in seed plants: Current understanding, applications, and perspectives. *Plant Commun.* 3, 100328. doi: 10.1016/j.jxplc.2022.100328
- Yang, F., Li, W., Jiang, N., Yu, H., Morohashi, K., Ouma, W. Z., et al. (2017). A maize gene regulatory network for phenolic metabolism. *Mol. Plant* 10, 498–515. doi: 10.1016/j.molp.2016.10.020
- Yang, F., Ouma, W. Z., Li, W., Doseff, A. I., and Grotewold, E. (2016). Establishing the Architecture of Plant Gene Regulatory Networks. *Methods Enzymol.* 576, 251–304. doi: 10.1016/bs.mie.2016.03.003
- Yang, Z. J., Wang, C. W., Xue, Y., Liu, X., Chen, S., Song, C. P., et al. (2019). Calcium-activated 14–3–3 proteins as a molecular switch in salt stress tolerance. *Nat. Commun.* 10, 12. doi: 10.1038/s41467-019-09181-2
- Yang, B., Xu, Y., Maxwell, A., Koh, W., Gong, P., and Zhang, C. (2018). MICRAT: a novel algorithm for inferring gene regulatory networks using time series gene expression data. *BMC Syst. Biol.* 12, 115. doi: 10.1186/s12918-018-0635-1
- Yang, X. D., Yang, S. P., Qi, H., Wang, T. P., Li, H., and Zhang, Z. D. (2020). PlaPPIsite: a comprehensive resource for plant protein-protein interaction sites. *BMC Plant Biol.* 20, 61. doi: 10.1186/s12870-020-2254-4
- Yao, M., Guan, M., Zhang, Z. Q., Zhang, Q. P., Cui, Y. X., Chen, H., et al. (2020). GWAS and co-expression network combination uncovers multigenes with close linkage effects on the oleic acid content accumulation in *Brassica napus*. *BMC Genomics* 21, 320. doi: 10.1186/s12864-020-6711-0
- Ye, Q., Zhu, F., Sun, F., Wang, T.-C., Wu, J., Liu, P., et al. (2022). Differentiation trajectories and biofunctions of symbiotic and un-symbiotic fate cells in root nodules of *Medicago truncatula*. *Mol. Plant* 15, 1852–1867. doi: 10.1016/j.molp.2022.10.019
- Yi, F., Gu, W., Chen, J., Song, N., Gao, X., Zhang, X. B., et al. (2019). High temporal-resolution transcriptome landscape of early maize seed development. *Plant Cell* 31, 974–992. doi: 10.1105/tpc.18.00961
- Yilmaz, A., Mejía-Guerra, M. K., Kurz, K., Liang, X., Welch, L., and Grotewold, E. (2011). AGRIS: the arabidopsis gene regulatory information server, an update. *Nucleic Acids Res.* 39, D1118–D1122. doi: 10.1093/nar/gkq1120
- Yu, C. P., Chen, S. C., Chang, Y. M., Liu, W. Y., Lin, H. H., Lin, J. J., et al. (2015). Transcriptome dynamics of developing maize leaves and genome-wide prediction of cis elements and their cognate transcription factors. *Proc. Natl. Acad. Sci. U.S.A.* 112, E2477–E2486. doi: 10.1073/pnas.1500605112
- Yuan, Q., and Duren, Z. (2024). Inferring gene regulatory networks from single-cell multiome data using atlas-scale external data. *Nat. Biotechnol.* Online ahead. doi: 10.1038/s41587-024-02182-7
- Yuan, Y., Huo, Q., Zhang, Z., Wang, Q., Wang, J., Chang, S., et al. (2024). Decoding the gene regulatory network of endosperm differentiation in maize. *Nat. Commun.* 15, 34. doi: 10.1038/s41467-023-44369-7
- Zander, M., Lewsey, M. G., Clark, N. M., Yin, L., Bartlett, A., Saldierna Guzmán, J. P., et al. (2020). Integrated multi-omics framework of the plant response to jasmonic acid. *Nat. Plants* 6, 290–302. doi: 10.1038/s41477-020-0605-7
- Zeng, W., Chen, X., Duren, Z., Wang, Y., Jiang, R., and Wong, W. H. (2019). DC3 is a method for deconvolution and coupled clustering from bulk and single-cell genomics data. *Nat. Commun.* 10, 4613. doi: 10.1038/s41467-019-12547-1

- Zhan, J., Thakare, D., Ma, C., Lloyd, A., Nixon, N. M., Arakaki, A. M., et al. (2015). RNA sequencing of laser-capture microdissected compartments of the maize kernel identifies regulatory modules associated with endosperm cell differentiation. *Plant Cell* 27, 513–531. doi: 10.1105/tpc.114.135657
- Zhang, Y., Han, E., Peng, Y., Wang, Y., Wang, Y., Geng, Z., et al. (2022a). Rice co-expression network analysis identifies gene modules associated with agronomic traits. *Plant Physiol.* 190, 1526–1542. doi: 10.1093/plphys/kiac339
- Zhang, L., He, C., Lai, Y., Wang, Y., Kang, L., Liu, A., et al. (2023b). Asymmetric gene expression and cell-type-specific regulatory networks in the root of bread wheat revealed by single-cell multiomics analysis. *Genome Biol.* 24, 65. doi: 10.1186/s13059-023-02908-x
- Zhang, Y., Li, Z., Liu, J., Zhang, Y., Ye, L., Peng, Y., et al. (2022b). Transposable elements orchestrate subgenome-convergent and -divergent transcription in common wheat. *Nat. Commun.* 13, 6940. doi: 10.1038/s41467-022-34290-w
- Zhang, Y., Li, Z., Zhang, Y., Lin, K., Peng, Y., Ye, L., et al. (2021). Evolutionary rewiring of the wheat transcriptional regulatory network by lineage-specific transposable elements. *Genome Res.* 31, 2276–2289. doi: 10.1101/gr.275658.121
- Zhang, F. Y., Liu, S. W., Li, L., Zuo, K. J., Zhao, L. X., and Zhang, L. D. (2016a). Genome-wide inference of protein-protein interaction networks identifies crosstalk in abscisic acid signaling. *Plant Physiol.* 171, 1511–1522. doi: 10.1104/pp.16.00057
- Zhang, S., Yang, W., Zhao, Q., Zhou, X., Jiang, L., Ma, S., et al. (2016b). Analysis of weighted co-regulatory networks in maize provides insights into new genes and regulatory mechanisms related to inositol phosphate metabolism. *BMC Genomics* 17, 129. doi: 10.1186/s12864-016-2476-x
- Zhang, B., Zhu, X., Chen, Z., Zhang, H., Huang, J., and Huang, J. (2023a). RiceTFtarget: A rice transcription factor–target prediction server based on coexpression and machine learning. *Plant Physiol.* 193, 190–194. doi: 10.1093/plphys/kiad332
- Zhao, N., Geng, Z., Zhao, G., Liu, J., An, Z., Zhang, H., et al. (2024). Integrated analysis of the transcriptome and metabolome reveals the molecular mechanism regulating cotton boll abscission under low light intensity. *BMC Plant Biol.* 24, 182. doi: 10.1186/s12870-024-04862-7
- Zhao, J. W., Lei, Y., Hong, J. W., Zheng, C. J., and Zhang, L. D. (2019). AraPPINet: an updated interactome for the analysis of hormone signaling crosstalk in *Arabidopsis thaliana*. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00870
- Zhao, T., Wu, H., Wang, X., Zhao, Y., Wang, L., Pan, J., et al. (2023). Integration of eQTL and machine learning to dissect causal genes with pleiotropic effects in genetic regulation networks of seed cotton yield. *Cell Rep.* 42, 113111. doi: 10.1016/j.celrep.2023.113111
- Zhong, R., Allen, J. D., Xiao, G., and Xie, Y. (2014). Ensemble-based network aggregation improves the accuracy of gene network reconstruction. *PLoS One* 9, e106319. doi: 10.1371/journal.pone.0106319
- Zhong, V., Archibald, B. N., and Brophy, J. A. N. (2023). Transcriptional and post-transcriptional controls for tuning gene expression in plants. *Curr. Opin. Plant Biol.* 71, 102315. doi: 10.1016/j.pbi.2022.102315
- Zhou, P., Li, Z., Magnusson, E., Gomez Cano, F., Crisp, P. A., Noshay, J. M., et al. (2020). Meta gene regulatory networks in maize highlight functionally relevant regulatory interactions. *Plant Cell* 32, 1377–1396. doi: 10.1105/tpc.20.00080
- Zhu, W., Miao, X., Qian, J., Chen, S., Jin, Q., Li, M., et al. (2023b). A translome-transcriptome multi-omics gene regulatory network reveals the complicated functional landscape of maize. *Genome Biol.* 24, 60. doi: 10.1186/s13059-023-02890-4
- Zhu, G., Wu, A., Xu, X. J., Xiao, P. P., Lu, L., Liu, J., et al. (2016). PPIM: A protein-protein interaction database for maize. *Plant Physiol.* 170, 618–626. doi: 10.1104/pp.15.01821
- Zhu, T., Zhou, X., You, Y., Wang, L., He, Z., and Chen, D. (2023a). cisDynet: An integrated platform for modeling gene-regulatory dynamics and networks. *iMeta* 2, e152. doi: 10.1002/imt2.152
- Zuin, J., Roth, G., Zhan, Y., Cramard, J., Redolfi, J., Piskadlo, E., et al. (2022). Nonlinear control of transcription through enhancer-promoter interactions. *Nature* 604, 571–577. doi: 10.1038/s41586-022-04570-y