



A multi-layered systems biology framework reveals dual-phased regulators and hormonal crosstalk underlying soybean cold tolerance

Hao-Yu Liu^{1,2} · Pei-Hsiu Kao³ · Supaporn Baiya⁴ · Chung-Feng Kao^{2,5}

Received: 18 July 2025 / Accepted: 20 October 2025 / Published online: 5 November 2025
© The Author(s) 2025

Abstract

Key message SNFE framework identifies 10 key CTgenes and reveals novel cold-tolerance mechanisms in soybean. **Abstract** Cold stress poses a significant threat to soybean (*Glycine max* (L.) Merr) productivity, during early developmental stages. Traditional approaches for identifying cold-responsive genes have been limited by size bias, pathway redundancy, and lack of integrative validation. To address these challenges, we developed a multi-layered systems biology framework, termed SNFE (systems and network-based feature engineering), to uncover key cold-tolerant genes (CTgenes) by leveraging both panomics and non-omics data in a network-informed context. The SNFE framework integrates five analytical layers: functional pathway enrichment, pathway crosstalk, co-functional network construction, network topology analysis, and experimental validation. From an initial pool of cold-responsive genes, SNFE identified 10 key CTgenes demonstrating high connectivity, regulatory importance, and consistent differential expression in short- and mid-term cold conditions. These genes were validated via independent transcriptomic datasets, Quantitative real-time PCR analysis, and hormone profiling. Notably, SNFE revealed novel regulatory mechanisms, including dual-timed transcription factors, ABA-JA hormone synergy in membrane stabilization, and convergence of abiotic and biotic stress signaling. A Sankey diagram and volcano plot further confirmed that most CTgenes reside at key regulatory nodes, linking upstream functions to downstream cold-tolerance pathways. SNFE is a reliable, efficient, and interpretable tool that not only improves prediction accuracy but also enables the discovery of novel biological insights. Its scalability and analytical depth make it a powerful platform for dissecting complex stress responses in crops. This framework provides a strategic foundation for molecular breeding; we also discuss the potential of multiplex “full gene packages” as a downstream engineering avenue to enhance cold resilience.

Keywords Multi-layered systems biology · SNFE framework · Panomics · Hormone–membrane crosstalk · Dual-phase regulators · Full gene packages

Abbreviations

CT	Cold tolerance	OnO	Omics and Non-omics
GO	Gene ontology	ABA	Abscisic acid
QTL	Quantitative trait locus	SNP	Single-nucleotide polymorphism
GWAS	Genome-wide association study	FOSCO	First-Order Statistic Correction
		QQ-plots	Quantile–Quantile Plots
		qRT-PCR	Quantitative real-time polymerase chain reaction

Communicated by Kinya Toriyama.

✉ Supaporn Baiya
supaporn.bai@ku.th

✉ Chung-Feng Kao
kaoc@nchu.edu.tw

¹ Plant Sciences Group, Wageningen University and Research, Wageningen, The Netherlands

² Department of Agronomy, College of Agriculture and Natural Resources, National Chung Hsing University, Taichung, Taiwan

³ Faculty of Science, University of Melbourne, Parkville, Australia

⁴ Department of Resource and Environment, Faculty of Science at Sriracha, Kasetsart University at Sriracha Campus, Chonburi, Thailand

⁵ Advanced Plant and Food Crop Biotechnology Center, National Chung Hsing University, Taichung, Taiwan

HPLC–MS/MS	High-performance liquid chromatography–tandem mass spectrometry
GSEA	Gene set enrichment analysis
ES	Enrichment score
BH	Benjamini–Hochberg
PPI	Protein–protein interactions
SA	Salicylic acid

Introduction

Soybean (*Glycine max* (L.) Merr), a subtropical leguminous plant native to southeastern Asia, has served as a staple crop for millennia due to its high protein content, low glycemic index, and a range of beneficial phytochemicals (Badole and Bodhankar 2012). Its remarkable adaptability allows it to thrive across a wide geographical range (Kim et al. 2012), spanning latitudes from 20–50° N to 10–40° S, making it a critical contributor to global food security (Leff et al. 2004). However, despite this adaptability, soybean remains highly sensitive to low temperatures, particularly during its early developmental stages, posing a significant challenge. Cold stress—whether chilling (0–15 °C) or freezing (<0 °C)—threatens soybean production, especially in regions outside its native range. As climate change induces more frequent and severe temperature fluctuations (Ray et al. 2019), the development of cold-tolerant soybean cultivars has become essential for maintaining yield stability.

Cold tolerance in plants is a complex, polygenic trait governed by numerous genes and environmental interactions (De La Torre et al. 2021). Soybean, in particular, is highly sensitive to temperature fluctuations, suffering various adverse effects under cold stress (Staniak et al. 2021). Chilling temperatures disrupt membrane fluidity and promote the accumulation of reactive oxygen species, while freezing conditions lead to extracellular ice formation, causing cellular dehydration and mechanical damage (Lissarre et al. 2010; Ruelland et al. 2009). These physiological disruptions significantly impair soybean's growth cycle, particularly during crucial stages, such as flowering and pod development (Kokubun 2011). For instance, temperatures below 15 °C can severely inhibit growth, and temperatures around 10 °C may completely prevent flowering, resulting in substantial yield losses (Mariola Staniak 2021). This vulnerability to cold stress highlights the importance of uncovering the genetic mechanisms underlying cold tolerance and incorporating these traits into breeding programs for improved resilience.

In response to cold stress, plants undergo a variety of physiological and molecular changes. Key components of cold tolerance include membrane stability, antioxidant activity, and phytohormone signaling (Raza et al. 2023), alongside the activation of cold-responsive genes and

transcription factors (Lissarre et al. 2010). Notably, the CBF/DREB1 gene family plays a central role in cold acclimation by regulating the expression of downstream genes involved in cold response (Lissarre et al. 2010; Thomashow 2010). However, the complex interactions among these pathways remain incompletely understood. While recent advances in genomics have identified several candidate genes linked to cold tolerance, these studies often fail to encompass the broader gene interaction networks critical for a comprehensive understanding of cold stress response (Thomashow 2010).

Traditional approaches for identifying cold tolerance genes, such as quantitative trait loci (QTL) mapping (Collard et al. 2005) and genome-wide association studies (GWAS) (Visscher et al. 2012), have contributed valuable insights but are limited by challenges. The polygenic nature of cold tolerance complicates the identification of critical genes, and these methods often yield false positives or fail to capture important gene–environment interactions (Hesketh et al. 1973). In contrast, systems biology frameworks now increasingly adopt panomics datasets (genomics, transcriptomics, proteomics, metabolomics, phenomics, and emerging single-cell/spatial omics) coupled with computational modeling to provide holistic insights into plant stress responses (Raza et al. 2025).

Pathway enrichment analysis serves as a powerful approach to uncover the molecular mechanisms underlying complex traits such as cold tolerance (Mariola Staniak 2021). By evaluating the overrepresentation of specific biological pathways among cold-tolerant response genes (CTgenes) identified through a feature engineering framework incorporating both omics and non-omics (OnO) data (Kao et al. 2022), researchers gain functional insights into how these pathways contribute to cold stress adaptation. Cold stress, for instance, activates a cascade of transcription factors that regulate stress-responsive pathways related to osmotic balance, antioxidant defenses, and membrane fluidity (Ohnishi et al. 2010). When integrated with network analysis, pathway enrichment provides a more comprehensive understanding of how different pathways and gene modules coordinate to alleviate cold stress (Zeng et al. 2022).

In addition to transcriptomic data, panomics approaches have proven particularly valuable for elucidating cold tolerance mechanisms, especially when integrated with modern breeding tools (Raza et al. 2024). Cold stress induces significant metabolic adjustments, including the accumulation of protective osmolytes, flavonoids, and stress hormones such as abscisic acid (ABA), which aid in maintaining cellular homeostasis (Chinnusamy et al. 2010). Incorporating metabolomic data into systems biology studies allows researchers to link these metabolic changes to specific gene networks involved in cold

tolerance, providing a more comprehensive understanding of the physiological mechanisms underlying cold stress adaptation.

The identification of cold tolerance genes can be further refined through network analysis, which models biological relationships as interconnected nodes (genes, proteins, pathways) and edges (interactions) (Yu et al. 2021). This approach enables the visualization of complex interactions that regulate cold tolerance, highlighting how genes and pathways communicate to coordinate cold stress responses. For instance, constructing co-functional gene networks (Kim et al. 2017) allows for the prioritization of key CTgenes, which can be targeted in breeding programs to enhance cold tolerance in soybean.

Traditional gene discovery methods, such as comparative transcriptome analysis and reverse genetics, have been foundational in identifying genes related to cold tolerance. However, recent advances in bioinformatics, including GWAS (Priyanatha et al. 2022; Zhang et al. 2015) and pathway-based analyses (Cirillo et al. 2017), have provided greater precision. Despite these advancements, an integrated framework that combines omics data with network and pathway analysis is still needed (Roychowdhury et al. 2023). Consistent with recent panomics-based integrative frameworks under cold stress (Raza et al. 2024, 2025), we implemented a systems and network-based feature engineering (SNFE) workflow. This multi-layered framework incorporates functional pathway enrichment, pathway crosstalk, co-functional network construction, and network topology analysis, enabling the prioritization of key CTgenes. We further extend the panomics perspective by integrating omics and non-omics (OnO) layers to decode complex cold-tolerance mechanisms in soybean. Using transcriptomics data (Yamasaki and Randall 2016) with a previously identified pool of CTgenes in soybean, we aim to identify key CTgenes contributing to enhanced cold resilience.

A comprehensive systems biology approach is crucial for elucidating the genetic and molecular mechanisms underlying cold tolerance. In this study, we employed SNFE framework to integrate prior biological knowledge with panomics data, specifically targeting the identification of key CTgenes in soybean (Fig. 1). These CTgenes were selected based on their strong associations with cold stress responses across multiple datasets, and their functional roles were further investigated through integrative pathway analysis (Kao et al. 2022). Additionally, single-nucleotide polymorphism (SNP) correction was applied to account for genetic variation in each gene (Cheng and Zhu 2021). This SNFE framework provides valuable insights into the molecular mechanisms of cold tolerance, establishing a foundation for future research and breeding programs focused on enhancing cold resilience in soybean.

Results

Dataset compilation and candidate genes selection (multiple OnO data layer)

We previously identified 170 CTgenes, comprising 44 short-term and 143 mid-term CTgenes, from a comprehensive dataset of 60,726 genes derived from integrated OnO data. This dataset provided enhanced biological insights and functional mechanisms relevant to cold-tolerance responses in soybean. Validation was conducted using RNA-seq data from soybean seedlings exposed to 4 °C for 1 and 24 h, revealing critical molecular pathways underlying cold tolerance. The validated CTgenes were subsequently utilized as input for the SNFE framework to identify key CTgenes. Detailed methods are available in Kao et al. (2022).

Statistical pathway enrichment analysis (functional pathway layer)

Gene-wise statistic and correction

The application of the modified first order statistic correction (FOSCO) method effectively addressed gene size bias in gene-wise statistical analyses using SNP data from the 180 K AXIOM® SoyaSNP array (Lee et al. 2015). As depicted in Fig. 2A, the raw gene-wise statistic scores exhibited a right-skewed distribution, indicating an overrepresentation of larger genes, which possess higher SNP counts and, consequently, inflated significance scores. To rectify this bias, modified FOSCO was applied to normalize the statistic scores relative to gene size. The corrected scores displayed a uniform distribution across genes with varying SNP counts (Fig. 2B), demonstrating the effectiveness of the adjustment in mitigating the disproportionate influence of gene size. To further investigate gene size effects, genes were classified into small and large categories based on their SNPs counts. Genes with SNP counts below the dataset median were classified as small, whereas those with SNP counts equal to or greater than the median was designated as large. This classification facilitated the assessment of bias correction across different gene size groups. The quantile–quantile (QQ) plots provided additional validation of the correction's effectiveness. While the raw statistic scores showed significant deviations from the expected uniform distribution (Fig. 2C), the corrected scores closely aligned with the expected distribution (Fig. 2D), indicating the robustness of the normalization process. These results confirm that the modified FOSCO method provides an unbiased evaluation of gene significance, thereby offering a reliable foundation for subsequent pathway enrichment and network analyses.

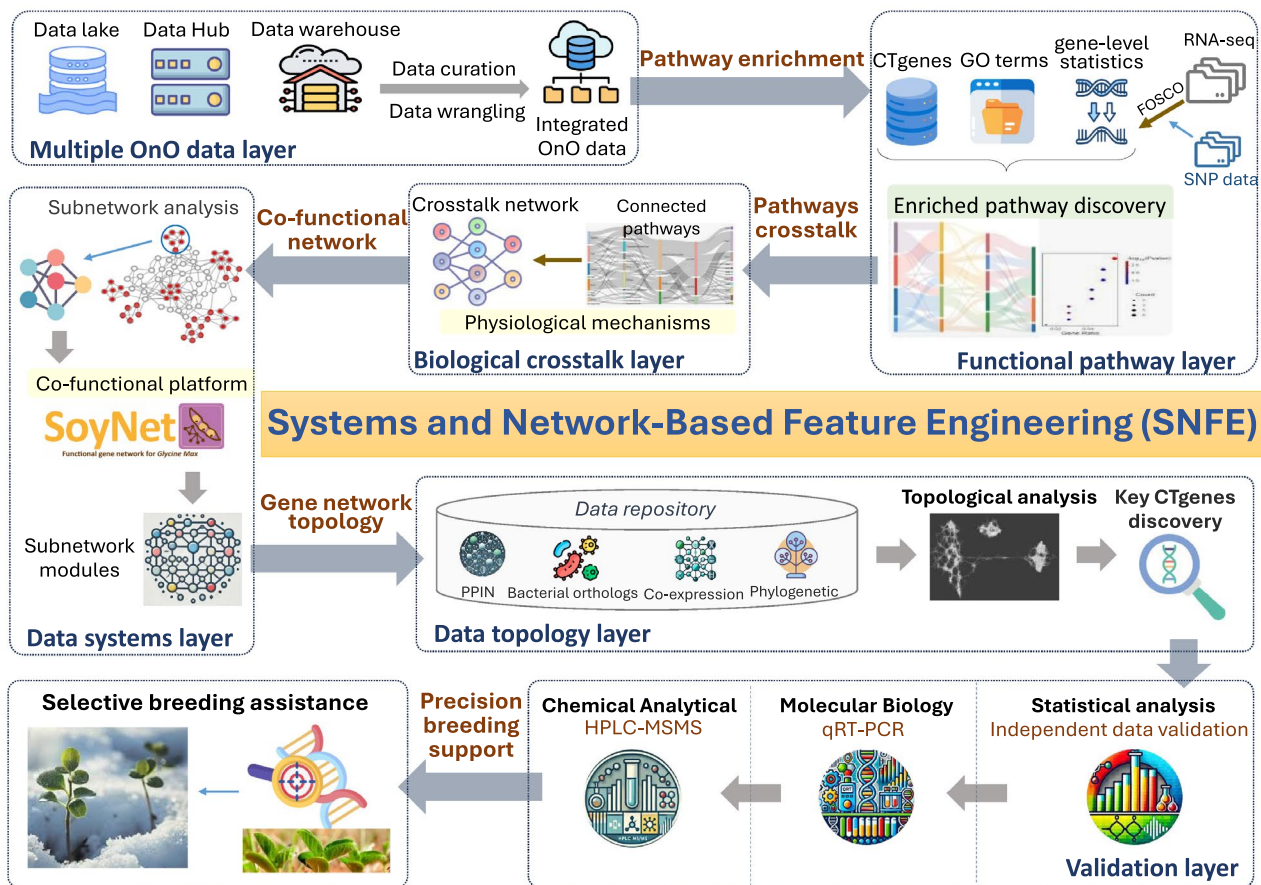


Fig. 1 Multi-layered systems biology framework for discovering key cold-tolerance genes in soybean using systems and network-based feature engineering (SNFE). This figure presents a systematic framework for identifying key cold-tolerant response genes (CTgenes) in soybean, comprising six interconnected layers that progressively refine and validate key CTgenes for functional relevance. In multiple OnO (omics and non-omics) data layer, a comprehensive gene pool (Kao et al. 2022), capturing cold-tolerant responses across diverse biological and environmental contexts, was used as the foundation for feature engineering. In functional pathway layer, pathway enrichment analyses were employed to identify pathways significantly associated with cold tolerance. Gene Ontology (GO) term (4,896 pathways) and external RNA-seq data (49,777 genes) (Yamasaki and Randall 2016) were utilized to access pathway enrichment, narrowing the gene pool to CTgenes linked to functional pathways. Gene-level statistics were calculated using gene expression p -values, corrected for gene size based on SNP data from 180K AXIOM® SoyaSNP array (Lee

et al. 2015). In biological crosstalk layer, interactions among enriched pathways were analyzed using crosstalk network analysis. Topological measures, such as Jaccard coefficients, were applied to identify strongly interconnected pathways and refine the CTgenes pool based on their roles in cold-tolerance biological processes. In data systems layer, co-functional network analysis was performed to construct subnetworks, identifying gene modules contributing to cold tolerance. In data topology layer, key CTgenes were prioritized based on their topological importance within subnetworks, using metrics, such as degree centrality and clustering coefficients. In validation layer, the identified key CTgenes were validated through molecular techniques (e.g., qRT-PCR), chemical methods (e.g., HPLC-MS/MS), and independent data validation. This multi-layered framework integrates multiple data sources, systematic biological analysis, and rigorous validation to discover and prioritize key CTgenes for precision breeding strategies aimed at enhancing cold tolerance in soybean

Pathway enrichment analysis

To identify pathways associated with cold tolerance, both competitive (Hypergeometric test, GSEA) and self-contained (SumStat, MaxMean) methods were employed, capitalizing on their complementary strengths to minimize biases and enhance pathway identification reliability. To ensure biological relevance, pathway selection was refined using thresholds based on CTgene count (≥ 5 for short-term

and ≥ 40 for mid-term) and CTgene ratio (≥ 0.05 for both phases) (Table S1). As illustrated in Fig. 3, a total of 22 pathways were identified as significantly enriched, including 13 associated with short-term (1 h) cold tolerance (Fig. 3A) and 11 with mid-term (24 h) cold tolerance (Fig. 3B). Among these, ‘response to water deprivation’ and ‘response to chitin’ pathways were consistently enriched across methods in both short-term and mid-term cold tolerance, underscoring their central role in stress adaptation. Pathways

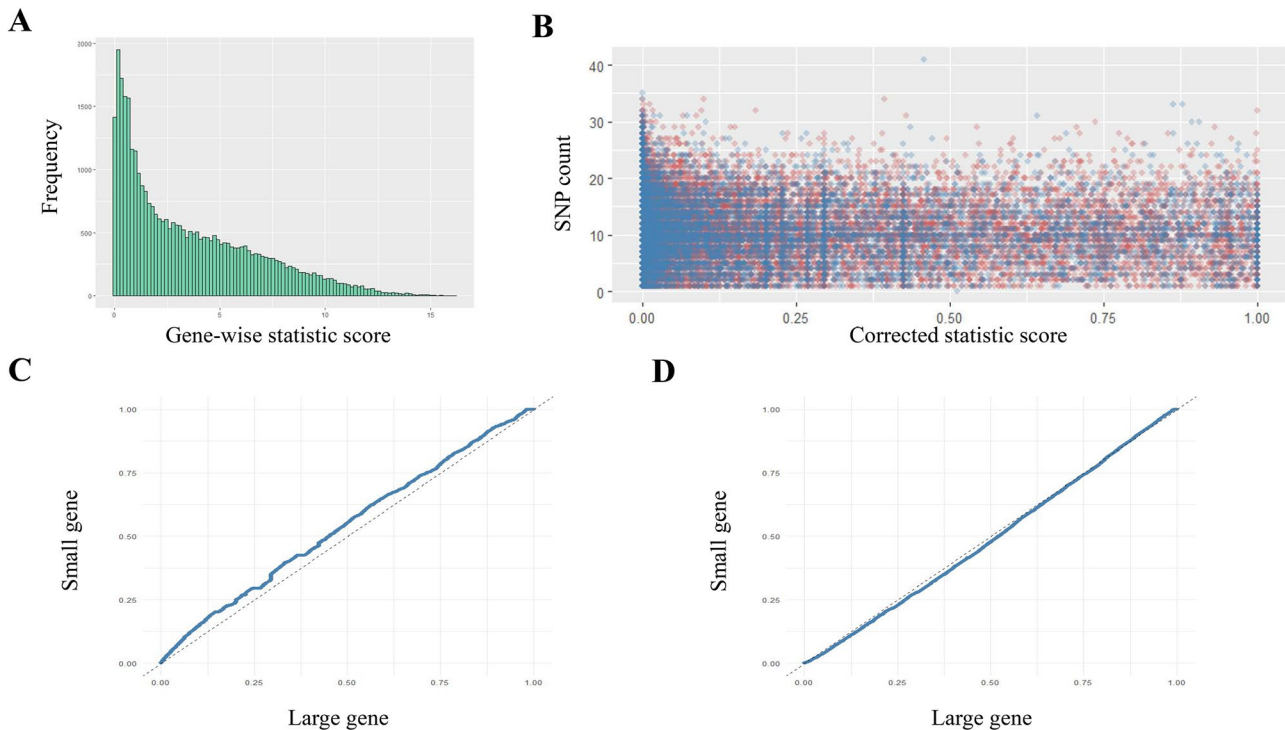


Fig. 2 Gene-wise statistic score correction. This figure illustrates the effect of gene size correction on gene-wise statistic scores, ensuring a balanced and unbiased analysis by mitigating gene size bias. **(A)** A bar chart displaying the distribution of raw gene-wise statistic scores, represented as negative logarithmic transformations. The right-skewed distribution indicates bias favoring genes with more SNPs. **(B)** A bar chart of gene size corrected statistic scores, showing a uniform distribution after adjustment for SNP counts per gene,

effectively reducing the overrepresentation of larger genes. Blue and red points denote genes in pathways and those not in pathways, respectively. **(C)** A QQ plot of raw statistic scores highlights deviations from the expected uniform distribution, reflecting gene size bias. **(D)** A QQ plot of gene size corrected statistic scores aligns with the expected uniform distribution, validating the effectiveness of the correction method

highlighted in red were consistently detected by multiple methods, emphasizing their relevance. Tables S1A and B further support these findings, presenting the top 0.2% of significantly enriched pathways with BH-corrected p -values ($p < 0.005$), confirming their strong association with cold-tolerance mechanisms. Consequently, 145 CTgenes (31 from short-term and 122 from mid-term) were selected from these pathways for pathway crosstalk analysis, which aims to identify interconnected pathways and further refine CTgenes selection by elucidating their roles within broader regulatory networks. Detailed results of pathway enrichment analysis are provided in Table S2.

Crosstalk of pathway with functional map (biological crosstalk layer)

To investigate the functional interconnectivity among enriched pathways, a crosstalk network analysis was conducted, mapping interactions between pathways associated with short-term (1 h) and mid-term (24 h) cold-tolerant responses (Fig. 4). The constructed network comprises 22 pathways connected by 102 edges, illustrating shared

functional roles in soybean cold stress adaptation. Among these, 13 pathways are primarily associated with short-term responses, while 11 pathways play a key role in mid-term adaptation. Notably, two pathways, ‘Response to water deprivation’ and ‘Response to chitin’, emerged as critical hub nodes, bridging both time frames and suggesting their central role in mediating cold tolerance across different stages.

To enhance biological interpretability, edges representing weak interactions (lower Jaccard coefficients) were filtered, retaining only the top 80% of connections for short-term pathways and 50% for mid-term pathways. This refinement yielded 61 significant interactions among the 22 pathways, providing a structured overview of cold tolerance mechanisms. Pathways related to jasmonic acid signaling, ABA response, and ethylene-mediated signaling exhibited strong interconnections, reflecting their coordinated roles in stress adaptation. Meanwhile, pathways associated with cold acclimation, photoprotection, and transcriptional regulation formed a distinct module, indicating their involvement in gene expression modulation under cold stress.

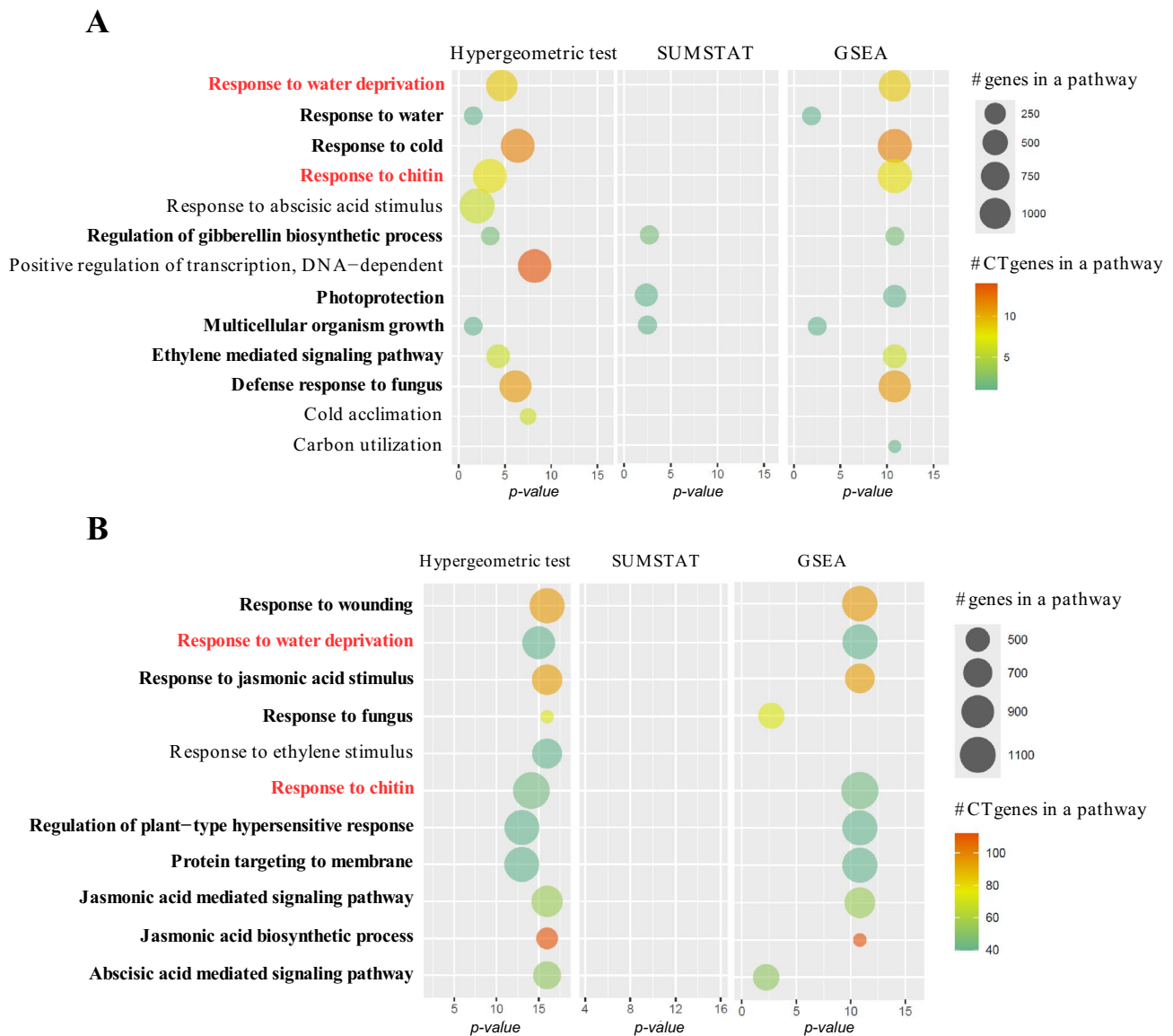


Fig. 3 Pathway enrichment analysis for short- and mid-term cold tolerance. This figure illustrates the outcomes of pathway enrichment analyses conducted using both competitive (hypergeometric test and GSEA) and self-contained (SumStat and MaxMean) methods for **(A)** short-term and **(B)** mid-term cold tolerance phases. Each node represents a pathway, with node size corresponding to the number of genes associated with the pathway. Node color indicates the number of CTgenes within the pathway, providing insight into their functional involvement. The x-axis displays the negative logarithmic transfor-

mation of p -values ($-\log p$ -value), reflecting statistical significance, while the y-axis shows the pathway annotations. Pathways highlighted in bold were consistently identified across at least two enrichment methods, underscoring their robustness. Pathways in red bold were consistently enriched across both short- and mid-term cold treatment periods, suggesting their critical role in mediating cold stress responses across different temporal phases. These results enhance the understanding of dynamic pathway regulation under cold stress conditions

This crosstalk network highlights functionally interconnected pathways that orchestrate cold stress responses, offering deeper insights into how soybean regulates physiological and molecular adaptations over time. The identification of key hub pathways underscores their potential as

targets for genetic improvement in cold-tolerant soybean breeding programs.

Co-functional networks analysis (data systems layer)

Building upon the results from the functional pathway and biological crosstalk layers, we utilized the 'Find Functional

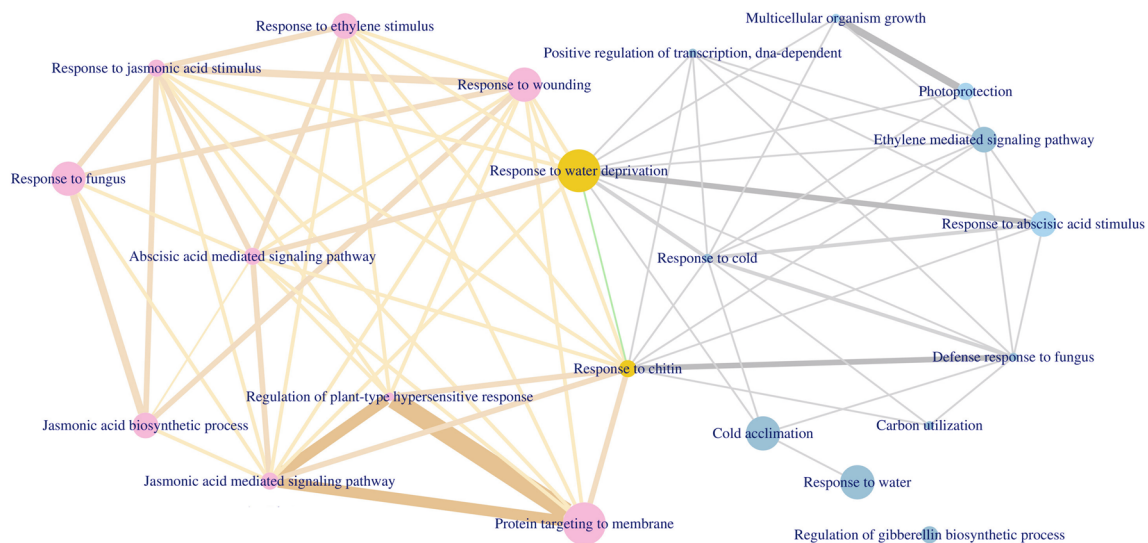


Fig. 4 Functional crosstalk network of enriched pathways under short- and mid-term cold tolerance in soybean. This figure illustrates a crosstalk network constructed from 22 significantly enriched pathways associated with short-term (1 h) and mid-term (24 h) cold stress responses. Nodes represent individual pathways, with node size proportional to the number of genes involved in each pathway. The network is topologically organized into three modules: blue nodes (right) represent short-term pathways primarily associated with cold acclimation, transcriptional regulation, and photoprotection; pink nodes (left) represent mid-term pathways enriched in jasmonic acid signaling, ABA response, and hypersensitive defense mechanisms;

and yellow nodes (center) indicate hub pathways enriched across both time points, highlighting their potential role in integrating immediate and sustained stress responses. Edges represent functional crosstalk between pathways, with edge width scaled by Jaccard coefficient values, indicating the degree of gene overlap between connected pathways. To improve network interpretability, only the top 80% of short-term and 50% of mid-term pathway interactions were retained. This crosstalk network reveals coordinated regulatory circuits underpinning temporal cold stress adaptation and suggests key transitions in signaling dynamics from early to prolonged exposure

Modules' algorithm from SoyNet (Kim et al. 2017), a soybean-specific co-functional database, to construct co-functional subnetworks and prioritize CTgenes based on their roles within these subnetworks to identify highly interconnected pathways associated with cold tolerance (Fig. 5). Pathways exhibiting strong functional associations during mid-term cold stress were grouped into subnetwork 1 (Fig. 5A), which includes 'Protein targeting to membrane', 'Jasmonic acid mediated signaling pathway', and 'Regulation of plant-type hypersensitive response'. Other mid-term pathways with lower interconnectivity were assigned to subnetwork 2 (Fig. 5B). Pathways enriched in the short-term cold response were clustered into subnetwork 3 (Fig. 5C). Notably, pathways, such as 'Response to water deprivation' and 'Response to chitin', which were consistently enriched across both short- and mid-term responses, were categorized into subnetwork 4 (Fig. 5D). Finally, subnetwork 5 (Fig. 5E) comprised pathways appearing in both short- and mid-term assessments, suggesting their sustained role in cold adaptation.

To refine candidate CTgenes within these subnetworks, we employed a data-driven approach (Valdés-Pérez 1999) integrating pathway interconnectivity and gene co-functionality. The classification resulted in 58, 122, 31, 68, and 145 CTgenes assigned to subnetworks 1 through 5, respectively. Specifically, subnetwork 1 (Fig. 5A) contained 58 CTgenes, primarily associated with membrane trafficking and jasmonic acid signaling, which are crucial for stress signal transduction. Subnetwork 2 (Fig. 5B) included 122 CTgenes, representing pathways involved in secondary metabolic processes and defense responses. Subnetwork 3 (Fig. 5C) featured 31 CTgenes, mainly linked to immediate cold stress response pathways, indicating their early-stage regulatory roles. Subnetwork 4 (Fig. 5D) contained 68 CTgenes, highlighting shared genes between short- and mid-term cold tolerance, reinforcing their importance in sustained stress adaptation. Finally, subnetwork 5 (Fig. 5E) included 145 CTgenes, integrating pathways across both time periods, suggesting their fundamental role in cold tolerance mechanisms. These resulting co-functional networks highlight key regulatory modules underlying soybean cold tolerance,

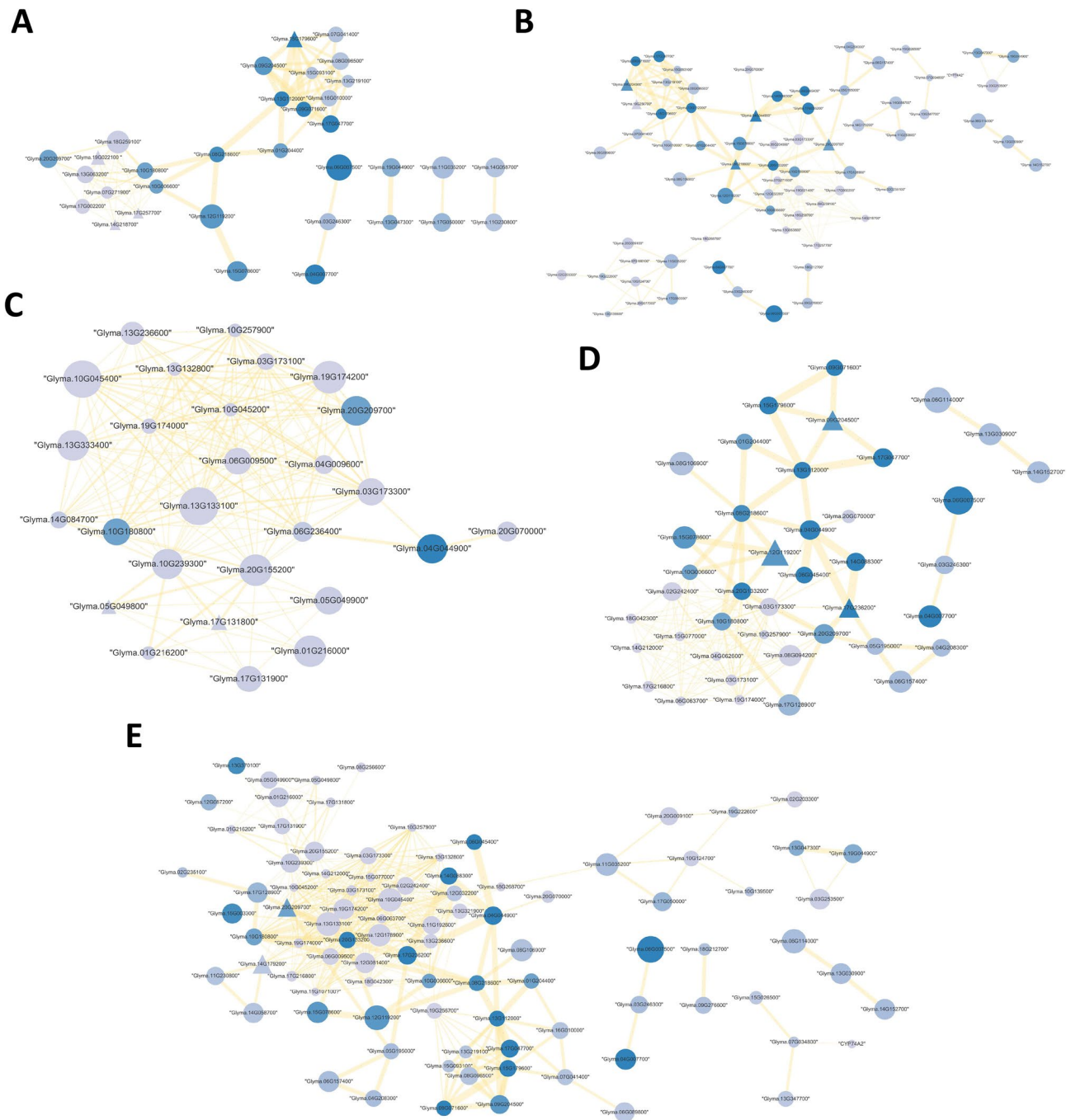


Fig. 5 Co-functional network analysis for key CTgenes discovery. This figure illustrates the co-functional network analysis used to identify key CTgenes based on pathway interconnectivity and functional relevance. Five subnetworks were constructed to classify cold tolerance-associated pathways: **(A)** Subnetwork 1, consisting of highly interconnected pathways during mid-term assessments; **(B)** Subnetwork 2, encompassing all other mid-term pathways with lower interconnectivity; **(C)** Subnetwork 3, representing short-term pathways responding to immediate cold stress; **(D)** Subnetwork 4, highlighting pathways consistently identified in both short- and mid-term cold tol-

erance phases; and **(E)** Subnetwork 5, integrating pathways appearing across both time periods, emphasizing their sustained role in cold adaptation. In the network visualization, darker nodes represent gene with higher expression levels, larger nodes indicate genes frequently occurring in enriched pathways, and triangular nodes mark the key CTgenes prioritized for their regulatory significance. Edge thickness reflects the frequency of gene co-occurrence within these pathways, indicating strong functional associations. The integration of pathway-level and network-level features provides a structured framework for prioritizing CTgenes critical to soybean cold tolerance

providing a structured framework for further experimental validation and breeding strategies.

Prioritization of key genes (data topology layer)

Following the co-functional network analysis, topology-based prioritization was applied to refine the identification of key CTgenes. Initially, 33, 70, 27, 41, and 92 genes were assigned to subnetworks 1 through 5, respectively, to assess their network significance, connectivity degree and clustering coefficients were calculated, revealing distinct connectivity patterns across the subnetworks. Specifically, connectivity degree ranged from 0 to 13 (average 5.03) in subnetwork 1, 0 to 16 (average 4.89) in subnetwork 2, 0 to 21 (average 4.88) in subnetwork 3, 0 to 18 (average 6.83) in subnetwork 4, and 0 to 32 (average 10.33) in subnetwork 5 (Table S3). To prioritize key CTgenes, genes with a connectivity degree near the median (50%) and a clustering coefficient above 0.5 were selected, refining the gene pool to 2, 5, 2, 1, and 2 key CTgenes from subnetworks 1 through 5, respectively. As a result, 10 key CTgenes were identified: *Glyma.17G131900*, *Glyma.05G049900*, *Glyma.15G179600*, *Glyma.09G204500*, *Glyma.08G218600*, *Glyma.19G256700*, *Glyma.20G209700*, *Glyma.04G044900*, *Glyma.17G236200*, and *Glyma.14G179200*.

Functional characterization of these 10 key CTgenes suggests diverse roles in cold stress responses (Fig. 6A–B). Specifically, *Glyma.09G204500*, *Glyma.05G049900*, and *Glyma.17G131900* function as transcription activators, while *Glyma.19G256700*, *Glyma.04G044900*, *Glyma.15G179600*, and *Glyma.17G236200* act as transcription repressors. Additionally, *Glyma.20G209700* and *Glyma.08G218600* belong to the transcription factor family (Consortium 2024; Tian et al. 2020), suggesting regulatory control over cold-responsive pathways. The remaining gene, *Glyma.14G179200*, currently lacks a well-defined function but exhibits strong network connectivity, warranting further investigation. Notably, several of these key CTgenes are functionally linked to the CBF regulatory pathway, a critical component of the plant cold response mechanism. Additionally, their association with jasmonate signaling pathways highlights their potential role in hormone-mediated cold stress adaptation. The integration of network topology metrics into the gene prioritization process provides a robust and systematic approach for uncovering key regulators of cold tolerance in soybean. For more detailed information, please see Table S4.

Validation for the key CTgenes (validation layer)

Statistical analysis

To evaluate the effectiveness and reliability of our SNFE framework in identifying key CTgenes, we conducted a comprehensive statistical validation using comparative expression analyses across different gene categories and cold treatment durations. As shown in Fig. 7, key CTgenes were compared against other CTgenes, intermediate genes, and remaining genes across short-term, mid-term, and combined cold response subnetworks.

In Fig. 7A, which represents the short-term subnetworks (subnetwork 3), the mean expression scores of key CTgenes were significantly higher than those of other gene categories during the 1 h cold stress period, indicating that these prioritized genes are highly responsive to early cold stimuli. This early responsiveness underscores the framework's sensitivity in capturing genes rapidly activated under chilling stress. By 24 h, key CTgenes exhibited significantly higher mean scores than intermediate and remaining genes, though the differences between key CTgenes and other CTgenes were less pronounced. Figure 7B, which focuses on mid-term subnetworks (subnetworks 1 and 2), shows that during the 24 h treatment, CTgenes consistently outperformed intermediate and remaining genes in terms of expression, while key CTgenes also displayed elevated mean scores. Although not significantly different from CTgenes, the consistent upregulation of key CTgenes supports the reliability of their prioritization in mid-term cold response. This further validates the framework's capacity to detect both early and sustained stress-responsive candidates. In Fig. 7C, which integrates subnetworks from both short- and mid-term responses (subnetworks 4 and 5), the key CTgenes again exhibited the highest mean scores, particularly during the mid-term period, followed by CTgenes and intermediate genes. This consistency across subnetworks and cold exposure durations affirms the robustness of our approach in selecting biologically relevant genes across temporal dynamics.

Together, these statistical analyses demonstrate that the key CTgenes identified through the SNFE framework exhibit significantly elevated expression under cold stress conditions, validating their central role in cold tolerance mechanisms and highlighting the framework's robustness in gene prioritization.

Molecular biology analysis

To validate the biological relevance of the 10 key CTgenes identified through the systems framework, we conducted transcriptional profiling using quantitative real-time polymerase chain reaction (qRT-PCR) under cold stress conditions. Soybean seedlings were subjected to cold treatment

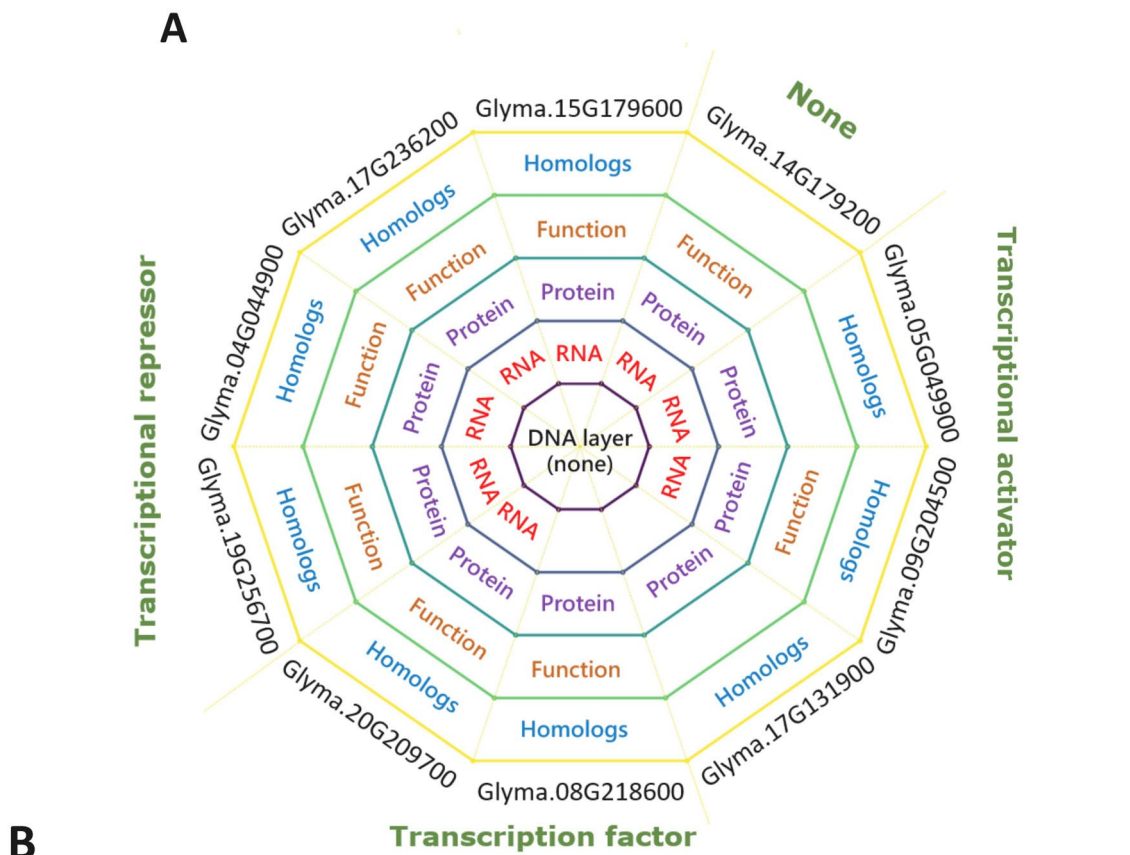


Fig. 6 Functional characterization of 10 key CTGenes identified through network topology analysis. **(A)** Sunburst chart illustrating the classification of the 10 key CTGenes involved in soybean cold-tolerance responses, highlighting their distinct functional roles. Genes are

categorized based on their regulatory function, including transcription activators, transcription repressors, and transcription factors, while others remain functionally undefined. **(B)** Summary of detailed information for each of the 10 key CTGenes

at 4 °C, and relative gene expression levels were assessed in both shoots and roots at six time points (1, 3, 6, 12, and 24 h) (Fig. 8). Among the ten key CTGenes, four genes—*Glyma.05G049900*, *Glyma.08G218600*, *Glyma.14G179200*, and *Glyma.17G131900*—exhibited dominant

expression in the shoot tissues. In contrast, the remaining six genes—*Glyma.04G044900*, *Glyma.09G204500*, *Glyma.15G179600*, *Glyma.17G236200*, *Glyma.19G256700*, and *Glyma.20G209700*—were more highly expressed in the roots.

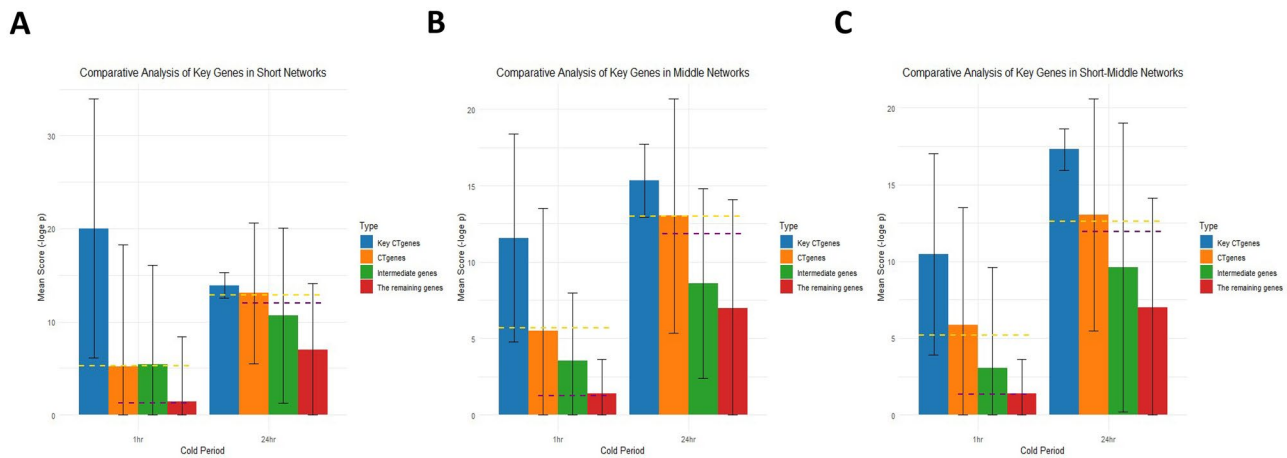


Fig. 7 Validation of key CTgenes through comparative gene expression analysis under cold stress conditions. This figure presents a comparative evaluation of the relative expression levels of 10 key CTgenes against other gene categories, including CTgenes, intermediate genes, and remaining genes, in soybean seedlings subjected to cold treatment. Panel (A) depicts results from short-term cold tolerance networks, (B) shows mid-term cold tolerance networks, and (C) illustrates combined short- and mid-term subnetworks. Colored bars

represent the mean expression scores ($-\log_{10} p$ -value) for each gene group, while error bars indicate standard deviations. Dotted lines denote thresholds used for statistical comparisons, with yellow indicating key CTgenes vs. other CTgenes, and purple representing comparisons against intermediate and remaining genes. The consistent upregulation of key CTgenes across conditions supports their central role in cold stress response and validates the gene prioritization strategy

In the shoots, several genes showed strong early induction. *Glyma.04G044900*, *Glyma.05G049900*, *Glyma.08G218600*, and *Glyma.20G209700* were significantly upregulated at 1 and 3 h post-treatment. Notably, *Glyma.05G049900* maintained elevated expression until 12 h, whereas *Glyma.20G209700* peaked at 1 h before declining. Expression of *Glyma.09G204500*, *Glyma.14G179200*, *Glyma.17G131900*, and *Glyma.17G236200* gradually increased, with maximum expression generally observed at 3 h. *Glyma.15G179600* exhibited its highest transcriptional level at 3 h. In contrast, *Glyma.19G256700* displayed overall downregulation in shoots.

In root tissues, six key CTgenes (*Glyma.04G044900*, *Glyma.08G218600*, *Glyma.09G204500*, *Glyma.15G179600*, *Glyma.17G236200*, and *Glyma.20G209700*) were strongly induced as early as 1 h post-treatment and remained significantly upregulated at 3 h. Of these, *Glyma.09G204500*, *Glyma.15G179600*, and *Glyma.17G236200* sustained elevated expression up to 6 h. Additionally, *Glyma.14G179200* and *Glyma.19G256700* reached their peak expression at 1 h and 6 h, respectively. In contrast, both *Glyma.05G049900* and *Glyma.17G131900* were downregulated in roots relative to control conditions.

These dynamic transcriptional responses across time points and tissues confirm the involvement of these 10 key CTgenes in early and sustained cold stress responses. The temporal patterns of expression, along with their organ-specific activation, support the effectiveness of our multi-layered systems biology framework in identifying biologically meaningful candidate genes for cold tolerance in soybean.

Plant hormones analysis

To further support the functional relevance of the identified key CTgenes, we quantified key plant hormones associated with abiotic stress signaling using High-performance liquid chromatography-tandem mass spectrometry (HPLC-MS/MS). Five plant hormones—SA, ABA, gibberellic acid (GA), kinetin (KI), and indole-3-acetic acid (IAA)—were analyzed in the shoots of soybean seedlings subjected to cold stress. Among the five hormones, only SA and ABA were detected at quantifiable levels under cold stress conditions, whereas GA₃, KI, and IAA were below detection limits, suggesting limited involvement of these hormones in early cold stress response in soybean seedlings.

As shown in Fig. 9A, SA concentrations remained unchanged relative to control levels from 1 to 6 h after cold treatment. However, a sharp and significant increase was observed at 12 h, with SA levels rising more than 34-fold compared to the control ($p < 0.05$), followed by a slightly reduced but still elevated level at 24 h (26-fold increase, $p < 0.05$). These findings indicate that SA plays a role in the late phase of the cold stress response. Similarly, ABA levels showed a significant increase at 12 h post-treatment (Fig. 9B), while no statistically significant differences were observed at earlier (1–6 h) or later (24 h) time points. This suggests a transient induction of ABA under cold conditions, potentially involved in the modulation of cold-responsive gene expression and stomatal regulation.

Together, these hormone profiling results align with the temporal expression patterns of several key CTgenes

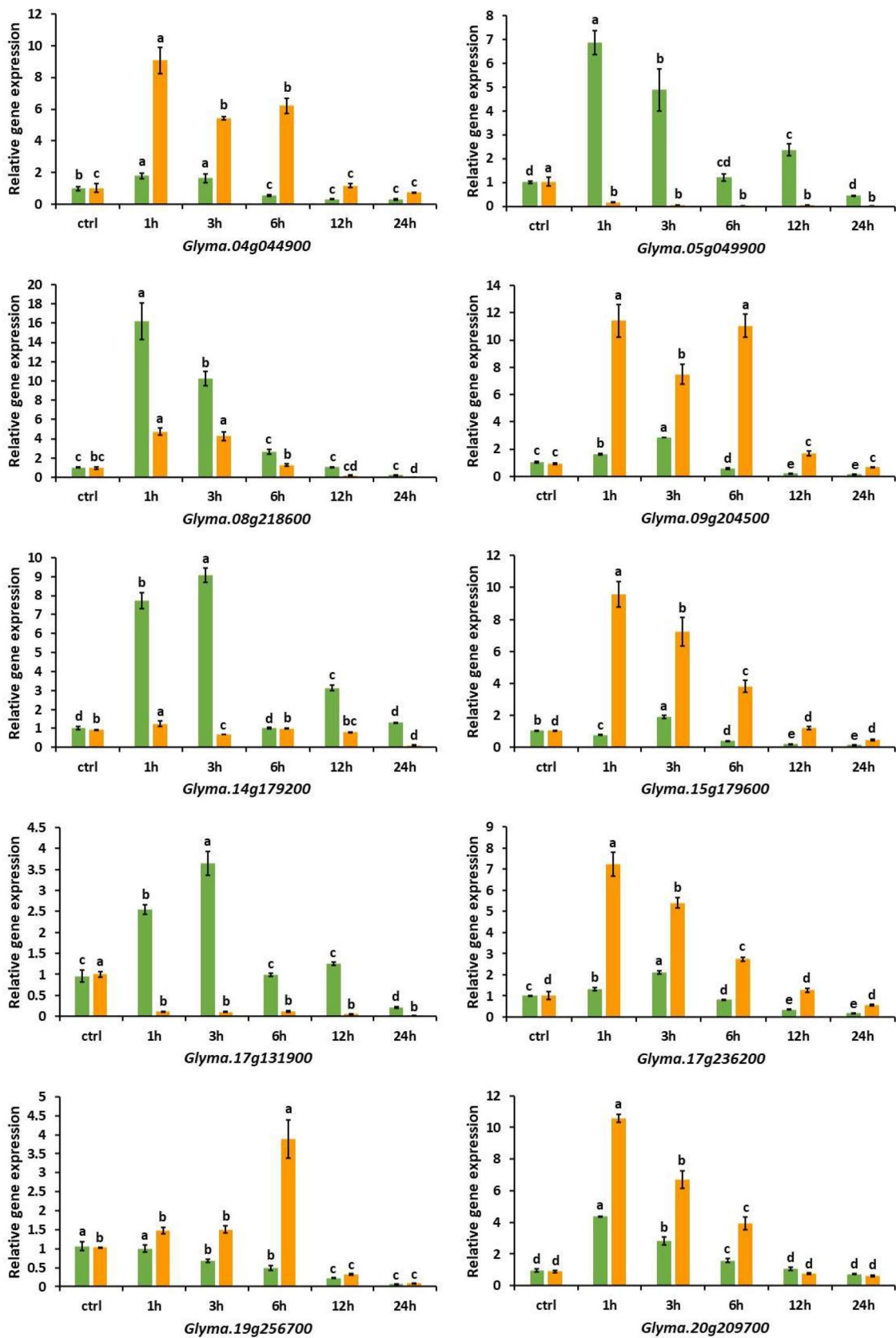


Fig. 8 Transcriptional profiling of key CTgenes in soybean seedlings under cold stress. qRT-PCR analysis of ten key CTgenes was performed to evaluate their transcriptional responses in soybean seedling shoots (green bars) and roots (orange bars) following exposure to cold stress. Gene-specific primers were used for quantification, and actin served as the internal reference gene. Expression levels were normalized to untreated control plants (ctrl). Error bars represent the standard error of the mean (\pm SE) from three biological replicates. Different lowercase letters above the bars denote statistically significant differences between time points, determined by ANOVA followed by Tukey's HSD test ($p < 0.05$). These results reveal distinct temporal and tissue-specific expression patterns across the ten key CTgenes, supporting their roles in cold stress adaptation

identified in this study—particularly those implicated in hormone-mediated stress signaling pathways—thereby validating the reliability and biological relevance of our integrative framework.

Discussion

Comparative insights with previous studies and recent panomics advances

Building upon our earlier work (Kao et al. 2022), which identified candidate CTgenes through statistical prioritization of integrated OnO features, the current study introduces a more mechanistically informed and network-driven framework, termed SNFE. While our previous study successfully stratified gene responses across short-, mid-, and long-term cold stress phases, it did not incorporate network-level topology or pathway-based interactions, limiting its capacity to resolve hierarchical regulators or stress signaling convergence points. In contrast, SNFE employs five interlinked analytical layers, including functional pathway enrichment, pathway crosstalk, co-functional network construction, network centrality analysis, and experimental validation, to systematically refine and validate CTgenes from a broad cold-responsive gene set. This architecture enabled the discovery of regulatory features not previously detected, such as dual-phased response timing (e.g., *Glyma.04G044900*), hormone–membrane coordination (notably ABA–JA synergy), and hub genes integrating abiotic and biotic stress cues.

Importantly, our approach aligns with and extends recent developments in panomics-based stress biology. Raza et al. (2024) emphasized the value of integrating transcriptomic, proteomic, metabolomic, and phenomic datasets to unravel multilayered plant responses under cold stress, particularly when such datasets are analyzed within pathway-centric or modular frameworks. Their findings demonstrated how cold-induced proteomic shifts modulate downstream hormonal and transcriptional cascades, reinforcing the need for multi-omics integration. Expanding

this perspective, Raza et al. (2025) proposed the concept of “full gene packages” and systems-level engineering for abiotic stress tolerance. They highlighted how network-informed selection of multi-target regulators can enable robust genetic interventions, especially under combined stress conditions.

By incorporating both Kao et al. (2022)'s foundational insights and the broader panomics paradigms introduced by Raza et al. (2024; 2025), the SNFE framework bridges the gap between data-driven discovery and functional interpretability. This not only advances predictive precision for CTgenes but also deepens our mechanistic understanding of cold tolerance networks, ultimately enhancing translational applications in soybean molecular breeding.

Pathway enrichment and crosstalk between short-term and mid-term responses

Functional pathway analysis revealed a clear dichotomy between early (short term) and later (mid-term) cold responses in soybean. In the first hours, upregulated genes were strongly enriched for cold acclimation processes, confirming that our early-response set captured canonical cold-responsive pathways. These genes were also enriched for negative regulation of gibberellin biosynthesis, indicating that rapid growth suppression is integral to the initial response. This pattern matches the physiological shift in plants that halts cellular expansion and reallocates energy to protective defenses (Achard et al. 2008). In parallel, we observed an overrepresentation of transcriptional control functions, consistent with the immediate activation of transcription factors that prime downstream protective programs.

By contrast, the mid-term response, occurring approximately 12 to 24 h after cold exposure, engages a broader set of stress and metabolic programs, consistent with extensive physiological adjustment under prolonged cold. Hormone signaling becomes prominent, with strong enrichment of ABA- and JA-related pathways. This concurrent activation points to hormone crosstalk during sustained exposure: ABA, classically tied to abiotic stress, and JA, often linked to defense, appear to act cooperatively rather than independently. Together, they coordinate downstream metabolic and defensive outputs that support continued cold tolerance.

Beyond hormone signaling, mid-term CT genes were enriched for defense and homeostatic processes, including categories linked to the hypersensitive response and respiratory burst (Table S1). These suggest recruitment of immune-associated circuits to shape ROS during prolonged cold. Consistently, hormone profiling confirmed a moderate late increase in SA (Fig. 9), a defense/HR-associated hormone, indicating activation of SA-dependent signals. Enrichment for ‘response to osmotic stress’ (Table S2), together with terms for protein targeting to

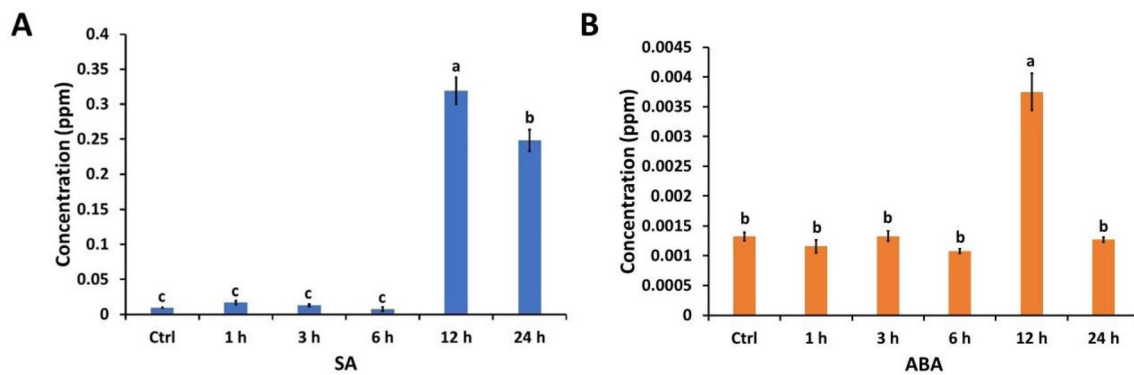


Fig. 9 Quantification of salicylic acid (SA) and abscisic acid (ABA) concentrations in soybean seedling shoots under cold stress. Hormone concentrations were measured using HPLC–MS/MS in multiple reaction monitoring mode across six time points (control, 1, 3, 6, 12, and 24 h) following cold stress treatment. Each bar represents the mean of five biological replicates. Hormone levels were quantified by compar-

ison to standard calibration curves and normalized using an internal standard (NAA). Values were adjusted by dividing by the measured NAA concentration and multiplying by the known amount of NAA added. Error bars represent \pm standard error. Different lowercase letters above columns denote statistically significant differences between time points ($p < 0.05$)

membranes and membrane fusion (Table S2), points to a proximal focus on sustaining membrane function under chilling. Because cold rigidifies membranes, active trafficking of transporters/aquaporins likely contributes to homeostasis (Afzal et al. 2016; Ruelland et al. 2009). Metabolic programs align with this view: unsaturated-fatty-acid biosynthesis is strongly enriched (Table S2)—a classic strategy to preserve membrane fluidity at low temperature (Zhao et al. 2021), and pathways for polyamines and amino acid catabolism (Table S2) support accumulation of compatible solutes that mitigate osmotic and oxidative damage (Cuevas et al. 2008; Szabados and Savouré 2010). Altogether, the pathway layer supports a temporal division of labor—early perception and transcriptional priming followed by a mid-term regime in which ABA-JA signaling, SA-modulated defense, and metabolic remodeling converge on membrane stabilization and homeostasis to sustain cold tolerance.

Crucially, the biological crosstalk layer allowed us to see how these pathways interconnect across time. Network mapping of enriched pathways and genes resolved distinct short-term and mid-term modules (Fig. 4 and Table S3), an early, relatively isolated module dominated by cold acclimation and transcription factor activity, and a highly interconnected mid-term module coupling hormone nodes (ABA, JA) to metabolic and defense processes. Points of intersection indicate temporal relays, early TFs likely regulate mid-term targets, and ABA and JA converge on shared downstream genes (Fig. 5), consistent with hormonal synergy. Several mid-term CTgenes span multiple categories, forming bridge nodes that connect, for example, JA signaling to oxidative stress outputs; notably, a JAZ-type repressor (*Glyma.15G179600*) links

the JA module to secondary metabolism (Wasternack and Hause 2013). Upstream inputs are also evident: enrichment of calcium-mediated signaling and MAPK components (Table S2) provides routes from early perception to hormonal control; cold-induced Ca^{2+} influx activates kinase cascades that elevate ABA and induce CBFs (Qian et al. 2024). In our network, a MAP-kinase node lies at the interface of ABA, JA, and HR pathways (Fig. 4 and Table S2), underscoring classical stress signaling as the organizer of hormone-defense crosstalk. Altogether, these features support a continuum in which early calcium/MAPK and transcriptional priming feed into mid-term ABA-JA control, thereby driving defense-like and metabolic adaptation for sustained cold tolerance.

Network topology highlights central hubs and temporal coordination

To refine the selection of candidate CTgenes, we constructed a co-functional network encompassing cold-responsive genes, referred to as the data systems layer, and analyzed its topology using quantitative network metrics (Table S3). Rather than detailing each metric, we emphasize the mechanism these rankings support: hub regulators organize the time structure of the response. An early hub (WRKY; *Glyma.08G218600*) aligns with rapid transcriptional priming; mid-term hubs, including a SCOF-like zinc finger (*Glyma.17G236200*) and a JAZ repressor (*Glyma.15G179600*) sit within hormone-rich modules; and dual-phase hubs, such as the ZAT10-like factor (*Glyma.04G044900*) and an R2R3-MYB (*Glyma.20G209700*), retain high connectivity from early into mid-term, indicating dual-phase function and serving

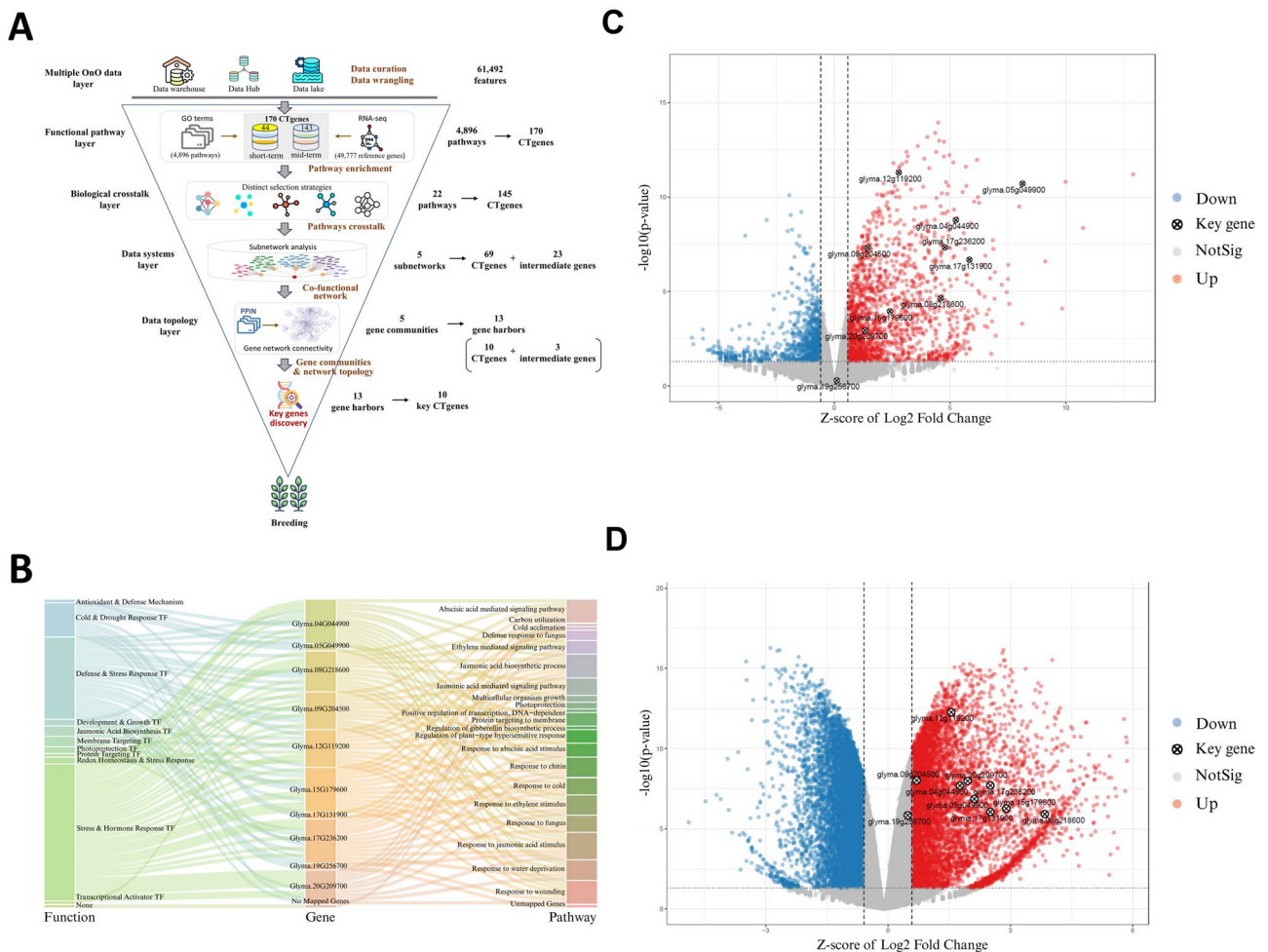


Fig. 10 Overview of the Systems and Network-Based Feature Engineering (SNFE) framework and multi-layered validation of cold tolerance genes. **(A)** The filtering process across different stages of the SNFE pipeline. In the Multiple OnO Data Layer, data from warehouse and lake sources were integrated through a data hub, processed via data wrangling and curation, resulting in 61,492 features. In the Functional Pathway Layer, 170 CTgenes (44 short-term and 143 mid-term) were analyzed against 4,896 Gene Ontology (GO) pathways using an RNA-seq reference dataset comprising 49,777 genes. Pathway enrichment narrowed the analysis to 170 CTgenes across 4,896 pathways. The Biological Crosstalk Layer further reduced this to 22 enriched pathways containing 145 CTgenes, leading to 5 functionally relevant subnetworks. In the Data Systems Layer, 69 CTgenes and 23 intermediate genes were retained. In the Data Topology Layer, these subnetworks were embedded into a co-functional network based on Protein–Protein Interaction Network (PPIN) data, resulting in the

identification of 5 gene communities containing 13 gene harbors (10 CTgenes and 3 intermediates). Network topology metrics ultimately refined this list to 10 key CTgenes for breeding applications. **(B)** Sankey diagram integrating transcriptional function categories (left), key CTgenes (center), and downstream enriched biological pathways (right). Node colors indicate transcription factor (TF) classes and biological processes. Flow lines denote functional associations based on gene annotation and pathway membership. **(C)** Volcano plot of gene expression under short-term (1 h) cold stress, based on \log_2 fold change and $-\log_{10} p$ -value. **(D)** Volcano plot of mid-term (24 h) cold stress response. Red and blue dots indicate significantly up- and down-regulated genes, respectively; black ‘x’ symbols highlight the 10 key CTgenes. Most key genes cluster in the upper right quadrant, confirming strong transcriptional induction under cold stress and supporting their regulatory relevance

as bridges from early perception (e.g., Ca^{2+} influx/ICE) to downstream acclimation. This pattern supports a continuum model in which bridging regulators hand off control from early perception to sustained hormone-metabolic adaptation.

To further demonstrate the predictive power and biological coherence of our multi-layered systems biology framework, we constructed a Sankey diagram (Fig. 10) to integrate three layers upstream transcriptional functions, the

ten CTgenes, and their downstream cold-responsive pathways. This view shows the CTgenes acting as mechanistic intermediates that couple high-level regulation to specific physiological programs. Nine of ten map to modules in hormone signaling (ABA, JA), membrane modification, or redox balance, processes central to cold tolerance, whereas *Glyma.19G256700* likely participates in under-annotated or non-canonical circuits. Thus, Fig. 10 underscores that our

systems-based pipeline prioritizes candidates that are not only statistically significant but also embedded in biologically meaningful positions within stress networks.

Together, Sankey integration (Fig. 10) and topology rank (Table S3) support a continuum model rather than a strict biphasic response: central, dual-phase hubs hand off control from early signaling to hormone-metabolic outputs, providing practical leverage points for engineering durable cold tolerance.

Experimental validation of key CTgenes

To validate the biological relevance of the 10 key CTgenes identified through our systems framework, we employed a combination of transcriptomic re-analysis, qRT-PCR, and hormone profiling. These validation approaches provided both *in silico* and *in vivo* confirmation of their involvement in soybean cold-stress responses (Fig. 10A).

We first conducted *in silico* validation using an independent transcriptome dataset by Yamasaki and Randall (2016) (Table S6), which profiled soybean gene expression responses to 4 °C cold treatment. Our analysis demonstrated that the 10 key CTgenes were consistently among the most significantly differentially expressed genes in this external dataset. (Fig. 7). Representative patterns align with their putative functions: the ZAT10-like *Glyma.04G044900* and the SCOF-like *Glyma.17G236200* are robustly induced under cold, whereas the JAZ homolog *Glyma.05G049900* is down-regulated, consistent with relief of JA repression (Table S6). These cross-dataset consistencies support the robustness of our prioritization.

To further demonstrate the predictive strength and biological coherence of our systems biology framework, we incorporated two complementary visualization strategies based on an independent RNA-seq dataset. First, we constructed a Sankey diagram (Fig. 10B) shows nine genes mapping to modules in hormone signaling (ABA/JA), membrane remodeling, or redox homeostasis, core processes in cold tolerance, while *Glyma.19G256700* remains unmapped, likely reflecting under-annotated or non-canonical circuitry. This integration underscores that our systems-based pipeline prioritizes candidates that are statistically strong and mechanistically embedded within stress response networks. Volcano plots under cold treatment position nine of ten CTgenes within the significantly up-regulated space, clearly separated from background genes (Fig. 10C–D).

Time-course qRT-PCR further corroborates *in vivo* responsiveness for all ten CTgenes (Fig. 8; Table S7). Early induction of the ZAT10-like *Glyma.04G044900* (peaking at ~ninefold in roots and subsequently declining toward baseline by 12–24 h) supports its role as a transition-phase activator in the cold response, whereas the SCOF-like *Glyma.17G236200* shows sustained induction at 12–24 h

consistent with a mid-term hub (Kim et al. 2001). The JAZ repressor *Glyma.15G179600* exhibited a progressive decline over time, indicating de-repression of JA signaling, which was accompanied by early induction of the JA-responsive WRKY transcription factor *Glyma.08G218600*, showing significant upregulation at 1 h, a secondary elevation at approximately 3 h, and a gradual decline from 6 to 24 h; the ABA-inducible bHLH *Glyma.19G256700* rises late, aligning with ABA-mediated responses.

Finally, we conducted hormone profiling to validate the physiological relevance of the hormonal pathways predicted from our network and crosstalk analyses. Targeted HPLC–MS/MS confirms a mid-term rise in ABA and a modest late increase in SA (Fig. 9A–B), matching pathway enrichments and crosstalk analyses. Together, these physiological trends support a model in which ABA–JA coordination underpins membrane/oxidative homeostasis, with SA contributing to defense-like ROS modulation during prolonged cold.

Together, these multilayered validation experiments confirm that the 10 key CTgenes identified are not only statistically robust but also functionally relevant. Their dynamic expression patterns and alignment with hormonal responses underscore their regulatory roles in orchestrating cold stress tolerance in soybean.

Novel insights into cold stress regulation in soybean

By integrating results across all analytical layers, our study uncovers several previously underexplored mechanisms that govern cold tolerance in soybean. These findings reveal a multifaceted and temporally coordinated response, involving transcriptional regulators, hormone signaling, membrane remodeling, and crosstalk between abiotic and biotic stress pathways.

Hormonal regulation of membrane adaptation

One of the most compelling insights from our analysis is the interplay between hormonal signaling and membrane remodeling during cold stress. In the mid-term, signatures converge on lipid desaturation and protein targeting to membranes, proximal effectors of membrane fluidity and integrity (Tables S1–S2). This suggests that maintaining membrane fluidity and integrity is a critical component of soybean's cold acclimation. Strikingly, this membrane remodeling appears to be coordinated by hormonal cues, particularly ABA and JA. Network and crosstalk analyses revealed that ABA- and JA-related pathways are functionally connected to lipid metabolism processes. Experimental data further support this model: cold exposure led to significant accumulation of ABA and modest increases in SA, with concurrent downregulation of a JA signaling

repressor (*Glyma.15G179600*, a JAZ1 homolog) and induction of an ABA-responsive bHLH transcription factor (*Glyma.19G256700*). These regulatory shifts likely activate downstream genes involved in lipid desaturation and membrane stabilization, aligning with studies in other species where ABA induces fatty acid desaturase expression and JA modulates lipid-derived secondary metabolites (Wang et al. 2023). This hormone-to-membrane coordination represents a novel and underexplored mechanism in cold tolerance. It suggests that cold-tolerant soybean genotypes may rely not only on canonical stress signaling pathways but also on hormone-mediated membrane remodeling to maintain cellular function under low temperatures. Notably, previous studies in tomato showed that JA can promote ABA biosynthesis under cold stress, enhancing tolerance (Ding et al. 2022). Consistently, our datasets show ABA-JA intertwining at both transcriptional and hormonal levels. The practical upshot is that breeders might target these hormone interactions, for instance, selecting for genotypes that more readily accumulate ABA and a small amount of SA/JA under cold, as this could lead to more robust membrane and metabolic adaptation.

Early and dual-phase transcriptional hubs orchestrate temporal coordination

Another important insight is the identification of early and dual-phase transcriptional regulators that anchor the transition from immediate cold perception to sustained defense (Qian et al. 2024). While the classical ICE-CBF-COR transcriptional cascade (Hwarari et al. 2022; Wang et al. 2025) was partially reflected in our data (enriched ‘response to cold’ pathway), none of the canonical CBF genes emerged as key CTgenes in our integrative analysis. Instead, two C2H2 zinc finger transcription factors, *Glyma.04G044900* and *Glyma.17G236200*, were identified as central network hubs. These TFs belong to the ZAT and SCOF families, which are well-documented in cold and oxidative stress responses (Kim et al. 2001; Wei et al. 2022). *ZAT10* (orthologous to *Glyma.04G044900*) functions downstream of CBFs to modulate osmotic and ROS stress tolerance, while SCOF-1 (related to *Glyma.17G236200*) acts as a transcriptional co-activator enhancing cold-responsive gene expression (Kim et al. 2001). Our data show that these TFs are not only strongly induced under cold stress but also maintain high network connectivity across both short- and mid-term phases, positioning them as critical dual-timed regulators. This dual-phase activity suggests that cold stress does not follow a rigid biphasic model but instead involves temporal handoffs between early and sustained regulators. Dual-timed TFs like *Glyma.04G044900* may integrate primary cold signals (e.g., calcium influx, ROS) and secondary hormone cues (ABA, JA), ensuring continuity of the response.

This insight could inform genetic strategies: a cultivar that strongly induces such dual-timed regulators might achieve a faster and longer-lasting acclimation to cold. Conversely, absence or weakness in these could break the chain between initial shock response and long-term adaptation.

Hormone-ROS crosstalk and recruitment of defense-like pathways

Our findings also indicate that cold tolerance in soybean involves components traditionally associated with biotic stress responses, particularly SA signaling and the HR (Song et al. 2023). Pathway enrichment analyses identified HR and ROS-related categories, and hormone profiling confirmed cold-induced accumulation of SA at later time points (12–24 h). This suggests that soybean employs a controlled oxidative burst under cold stress, likely mediated by SA and JA signaling, to activate downstream defenses. Supporting this, *Glyma.08G218600*, a WRKY transcription factor with known roles in integrating ROS and SA signaling, was identified as a cold-induced gene and a mid-level network hub. WRKYs, especially those in the GmWRKY40/85 family, have been implicated in both pathogen defense and abiotic stress responses. The concurrent upregulation of WRKYs, repression of JA repressors (JAZ1), and enrichment of SA/HR pathways suggests that cold tolerance leverages a low-level immune response as part of the acclimation strategy (Hu et al. 2013). This integrative mechanism could induce protective proteins like pathogenesis-related proteins and heat shock proteins, which function not only in pathogen defense but also in cellular stabilization under stress (Hu et al. 2013; Sadura and Janeczko 2024). These findings underscore the concept that abiotic and biotic stress pathways are not functionally distinct but rather converge at key regulatory nodes to maintain cellular integrity. The identification of specific regulators—zinc finger TFs, WRKYs, JAZ repressors, and ABA/SA-inducible genes—adds a new layer of mechanistic understanding to the cold stress response.

Together, these insights reframe soybean cold tolerance as a complex systems-level trait involving membrane remodeling, hormonal coordination, and transcriptional reprogramming across time. Our data expand the classical model of cold tolerance by highlighting JA/SA involvement, dual-timed transcriptional hubs, and membrane-hormone crosstalk. These findings not only provide new targets for breeding and genetic engineering but also illustrate how integrated multi-omics and network analysis can uncover hidden layers of regulation in plant stress responses.

Toward full gene package engineering for cold-tolerant soybean

The multi-layered nature of plant responses to cold stress involves a complex orchestration of signaling, transcriptional reprogramming, membrane remodeling, and metabolic adjustments. Traditional single-gene approaches may not fully capture the synergistic effects required for robust cold tolerance, particularly under fluctuating or combined environmental stresses. In this context, our identification of 10 key CTgenes enables the rational design of full gene packages, combinatorial gene modules that coordinate multiple tolerance pathways (Table S4).

The full gene package concept, as proposed by Raza et al. (2025), emphasizes the engineering of synergistic gene sets that target distinct but interconnected biological modules. Unlike additive transgene stacking, this approach prioritizes regulatory convergence and functional integration, enabling enhanced abiotic stress resilience through coordinated gene expression and pathway activation. It is particularly suited for crops like soybean, where cold stress tolerance is governed by a constellation of early-warning sensors, hormone signaling components, membrane stabilizers, and transcriptional regulators.

Based on our SNFE analysis, we propose several gene package combinations for future engineering of cold-tolerant soybean lines: (1) Priming-to-sustained response package: combining *Glyma.04G044900* (ZAT10-like, early priming) with *Glyma.17G236200* (SCOF-like, mid-term hub) to maintain early stress perception and extended defense activation. (2) Hormone-membrane package: pairing *Glyma.19G256700* (ABA-bHLH) with *Glyma.04G044900* to coordinate hormone-driven lipid desaturation and membrane stabilization. (3) ROS/defense-bridge package: combining *Glyma.08G218600* (WRKY) with *Glyma.20G209700* (R2R3-MYB) to stabilize oxidative signaling and integrate abiotic–biotic responses. (4) Negative regulator modulation: CRISPR-based attenuation of *Glyma.15G179600* (JAZ1) to prolong JA pathway activity, coupled with hormone priming strategies (methyl jasmonate or ABA).

Collectively, the full gene package strategy integrates systems biology, functional genomics, and network-guided design to move beyond traditional breeding. It offers a blueprint for next-generation soybean improvement pipelines, compatible with CRISPR multiplexing, synthetic regulatory elements, and panomics-informed predictive breeding. By combining mechanistic depth with translational flexibility, this approach may help realize climate-resilient agriculture in the face of escalating abiotic stress challenges.

Conclusion

This study introduces a robust and efficient multi-layered systems biology framework, SNFE, designed to precisely identify key CTgenes in soybean. By integrating OnO data across functional pathway enrichment, pathway crosstalk, co-functional network modeling, and topological prioritization, SNFE consistently pinpointed biologically relevant CTgenes. These candidates were consistently validated through independent transcriptomic datasets, qRT-PCR assays, and hormone profiling, confirming the framework's reliability and reproducibility. Importantly, SNFE uncovered novel regulatory mechanisms, such as dual-timed transcription factors, hormone–membrane coordination, and abiotic–biotic signaling integration, offering deeper insights into cold adaptation. The modular and scalable nature of SNFE supports its application to other complex traits, providing a powerful tool for precision breeding and stress-resilient crop development under climate variability.

Materials and methods

This study employed a multi-layered systems biology framework, referred to as SNFE, to discover and validate key CTgenes in soybean. The SNFE pipeline comprises six interconnected layers, each leveraging OnO data, statistical approaches, and computational tools to progressively refine and validate the key CTgenes. Figure 1 illustrates the framework.

Dataset compilation and candidate genes selection (multiple OnO data layer)

We utilized a set of 170 CTgenes prioritized from a foundational gene pool of 60,726 genes extracted from an integrated OnO dataset compiled in our previous study (Kao et al. 2022). The dataset was sourced from data lake, data hub, and data warehouse platforms to ensure comprehensive data collection and management. This dataset was constructed using a systems biology framework that integrated panomics (i.e., multi-omics) data including genomics, transcriptomics, proteomics, and metabolomics with non-omics data, such as physiological, phenotypic, environmental, pathological, and demographic attributes. The inclusion of homologous gene data from model plants further enriched that dataset, capturing prior knowledge across DNA, RNA, protein, functional, and ecological layers.

Feature selection and fitness evaluation methods were employed to refine the dataset, ensuring that the foundational gene pool of 60,726 genes was biologically relevant to soybean cold-tolerant responses. From this gene pool,

170 CTgenes were prioritized based on their potential functional importance in cold stress responses. This integrated approach combines the strengths of multi-omics data with non-omics contextual information, allowing for comprehensive biological insights and robust candidate gene selection. Such a framework provides a reliable foundation for downstream analyses aimed at uncovering molecular mechanisms underlying cold tolerance in soybean (see Table S5).

Statistical pathway enrichment analysis (functional pathway layer)

GO annotation

To refine the candidate genes for cold tolerance, pathway enrichment analyses were conducted using GO annotations, encompassing 4,896 pathways categorized into biological process (2,542 pathways), molecular function (1,898 pathways), and cellular component (456 pathways). Only pathways containing between 6 and 1,500 genes, along with at least one CTgene, were retained, resulting in 442 pathways (100 pathways associated with short-term cold stress and 342 pathways linked to mid-term stress conditions) relevant to cold-tolerant responses. This filtering ensured relevance to the study's focus and enhanced pathway reliability for further analysis.

Validation sample

For validation, an external RNA-seq dataset of 49,777 genes (Table S6), derived from cold-treated soybean leaves (Yamasaki and Randall 2016), was used. Soybean seedlings (cv. Williams 82) were subjected to 4°C for two days, with transcript data collected at 0, 1, and 24 h post-treatment. Gene expression levels were quantified using the 'edgeR' package in R, and alignment to the latest gene annotations from SoyBase (<https://www.soybase.org/>) resulted in a refined dataset of 42,576 genes, excluding 7,201 non-aligned genes.

Gene-wise statistic and correction

To address gene size bias in gene-level statistics, we applied a modified version of the FOSCO framework (Mirina et al. 2012). In this study, fold change (FC) was used instead of p -values to evaluate gene significance, offering a distinct perspective that emphasizes the biological impact of expression changes under specific conditions, such as cold tolerance. Unlike p -values, which focus on the likelihood

of observing changes by chance, FC highlights the magnitude of change, providing a direct measure of biological relevance.

Gene size bias arises because larger genes, with more SNPs, are more likely to show statistical associations (e.g., extreme p -values or FC values). This bias can lead to the overrepresentation of large genes in downstream analyses. By adjusting FCs for gene size, the modified FOSCO ensures that genes with extreme FC values are not disproportionately favored, resulting in a more balanced prioritization of candidate genes. Incorporating FC into FOSCO framework aligns statistical corrections with biological priorities, enhancing the biological relevance of gene rankings. The modified FOSCO algorithm is as follows:

1. Gene-level statistics calculation: FCs were converted into scores using a logarithmic transformation. Specifically, $S_{FC} = 10^{-\ln(\text{FC})}$ if $\text{FC} > 1$, and $S_{FC} = 10^{-\ln(\frac{1}{\text{FC}})}$ otherwise.
2. Gene size adjustment: The statistical scores were corrected for gene size using the number of SNP markers per gene, extracted from the 180K AXIOM® SoyaSNP array (Lee et al. 2015). The adjusted gene-level statistic was calculated as $S_{FC}^{\text{adj}} = 1 - (1 - S_{FC})^M$, where M represents the number of SNP markers within a gene, and λ is a tuning parameter empirically determined.
3. Optimal λ determination: The optimal λ value was determined by minimizing the absolute value of the correlation coefficient between S_{FC}^{adj} and M , obtained through a grid search. This ensures that the adjusted statistic is independent of gene size, effectively removing bias.

The modified FOSCO adjusts FCs by accounting for gene size, reducing the bias favoring larger genes and enabling unbiased statistical evaluation of CTgenes. By prioritizing genes based on biologically meaningful metrics and correcting for confounding factors such as gene size, this approach ensures robust and reliable pathway analysis. The corrected scores provide a more equitable basis for ranking candidate genes, facilitating downstream analyses of pathways and networks associated with cold tolerance in soybean.

Pathway enrichment analysis

Pathway enrichment analysis, a pivotal tool in genomics research, identifies biological pathways significantly associated with a specific phenotype or process by analyzing differential gene expression within gene sets (Goeman and Buhlmann 2007). This study employed both competitive and self-contained statistical methods (Ackermann and Strimmer 2009) to pinpoint pathways enriched with CTgenes, ensuring robustness and minimizing biases. These methods include the

hypergeometric test, gene set enrichment analysis (GSEA), SumStat, and MaxMean, each offering unique strengths in evaluating pathway significance.

Competitive methods for pathway enrichment analysis

Competitive methods, including the hypergeometric test and GSEA, were used to identify pathways significantly enriched with gene sets, such as CTgenes, by comparing their observed enrichment against a background dataset. These methods assess the statistical significance of the overlap between a given gene set and pathway-associated genes, enabling the identification of pathways relevant to the studied phenotype.

The hypergeometric test calculates the probability of observing an overlap between CTgenes and genes in a specific pathway, adjusting for the size of the pathway and the genome. This approach determines whether the observed overlap is greater than expected by chance. The p -value for the test is computed using the following formula:

$$p \text{ value} = \sum_{x=g}^s \frac{\binom{S}{x} \binom{L-S}{M-x}}{\binom{L}{M}}$$

where L is the total number of genes in the genome, S is the total number of genes in the pathway, M is the number of CTgenes, and x is the observed overlap between CTgenes and pathway genes. The hypergeometric test is particularly effective for identifying pathways enriched with a specific gene set by quantifying how likely the observed overlap is under random conditions.

GSEA evaluates whether genes from a specific pathway are predominantly located at the top or bottom of a ranked gene list, which is typically ordered by a metric such as FC or statistical significance. The enrichment score (ES) is calculated as the gene list is traversed. The ES increases when a pathway gene is encountered, proportional to its ranking, and decreases otherwise. The maximum ES, denoted as $ES(S)$, represents the enrichment level of the pathway. The p -value is derived by comparing $ES(S)$ with a null distribution generated by permuting the ranked gene list. The ES is defined as:

$$ES(S) = \max_{1 \leq i \leq N} \left\{ \sum_{g_j \in S, j \leq i} \frac{|r_j|^m}{N_R} - \sum_{g_j \notin S, j \leq i} \frac{1}{N - N_H} \right\}$$

Where N represents the total number of genes in the ranked list, i is the position of the gene in the ranked list, j

is the preceding position, r_j is the score of the gene at position j , $N^R = \sum_{g_j \in S} |r_j|^m$ is the sum of weighted scores for all genes in the pathway, and N_H denotes the number of genes in the pathway. These competitive methods complement each other by addressing both absolute enrichment (hypergeometric) and ranked enrichment (GSEA), making them robust tools for pathway analysis.

Self-contained methods for pathway enrichment analysis

Self-contained methods, such as SumStat and MaxMean, compare observed gene sets with randomly permuted sets to assess differential expression (Rahmatallah et al. 2012). These methods evaluate pathways independently without relying on a background dataset.

SumStat calculates the ES for each pathway by summing the statistical measures across gene sets containing CTgenes:

$$ES_{\text{SumStat}} = \sum_{i=1}^S t_i$$

where t_i represents the statistic for each gene in the pathway. The ES is compared against a null distribution generated by permuting gene sets 10,000 times to derive p -values (Subramanian et al. 2005).

MaxMean detects subtle gene expression changes by first converting gene statistics to z-scores (Efron and Tibshirani 2002; Ren et al. 2017). The method computes two scores, S_+ and S_- , representing positive and negative contributions, respectively:

$$S_+ = \frac{1}{n_s} \sum_{i \in S} z_i \cdot I\{z_i > 0\} \text{ and } S_- = -\frac{1}{n_s} \sum_{i \in S} z_i \cdot I\{z_i < 0\}$$

The final score, S , is determined as $S = \max(S_+, S_-)$. Randomized gene sets are generated to calculate a null distribution, and p -values are computed by comparing observed scores to permuted scores.

By leveraging these complementary approaches, this study minimizes biases, enhances the reliability of enriched pathway identification, and enables the detection of both strong and subtle signals, which is critical for understanding the complex regulatory networks underlying cold tolerance in soybean. To control type I errors, the Benjamini–Hochberg (BH) correction (Benjamini et al. 2001) was applied to manage the false discovery rate, ensuring that statistically significant results are reliable and not due to random chance.

Crosstalk network analysis (biological crosstalk layer)

Crosstalk network analysis was employed to investigate interactions among enriched pathways (Ballouz et al. 2015), providing critical insights into their functional relationships and shared roles in cold tolerance. The underlying hypothesis is that pathways involved in cold tolerance function as part of an interconnected network that coordinates complex physiological and molecular responses rather than acting in isolation.

In this analysis, pathways were represented as nodes, and edges between nodes were defined by Jaccard coefficients (Survarachakan et al. 2022), which quantify the overlap between pathway gene sets. The Jaccard coefficient, calculated as $J(A, B) = \frac{|A \cap B|}{|A \cup B|}$, where A and B represent the sets of genes associated with two pathways, was computed using the ‘igraph’ package in R (Csárdi et al. 2023). This metric provides a standardized measure of pathway similarity ranged from 0 to 1, with higher values indicating greater overlap and stronger functional connections between pathways. Strongly interconnected pathways were identified. The metric enabled the identification of hub pathways that play central roles in cold tolerance and facilitated the prioritization of CTgenes associated with these pathways. By identifying pathways with significant overlap, this analysis reveals synergistic interactions and co-regulatory mechanisms that may enhance the plant's ability to respond to cold stress. Unlike isolated pathway analyses, crosstalk network analysis provides a comprehensive view of the broader network context, emphasizing the importance of pathway interconnectivity in the cold-tolerance response.

Subnetworks or gene modules construction (data systems layer)

Data source

Gene interaction data were obtained from the SoyNet database (Kim et al. 2017), a curated resource that integrates experimental and predicted gene relationships specific to soybean. This database ensures that the analysis is biologically relevant and tailored to soybean-specific traits, reducing noise and enhancing the reliability of network-based insights.

Co-functional network analysis

To identify subnetworks or gene modules associated with cold tolerance, a co-functional network analysis was conducted using the SoyNet database (Kim et al. 2017).

In the constructed network, genes were represented as nodes, and edges between nodes denoted functional relationships, such as co-expression, shared pathways, or protein–protein interactions. The weight of each edge reflected the confidence level of the interaction, as provided by the SoyNet database (Kim et al. 2017).

Co-functional networks capture the relationships among genes that contribute to shared biological processes, providing a systems-level understanding of gene functionality under cold stress. Subnetworks, representing gene modules, were extracted using clustering algorithms to identify groups of genes with strong functional interactions. These modules are hypothesized to act as functional units that mediate complex traits like cold tolerance. The hypothesis underlying this analysis is that genes function collaboratively within interconnected modules rather than in isolation. By analyzing these networks, co-regulated or co-expressed genes involved in shared pathways can be identified, providing insights into their collective roles in adapting to cold stress. For network visualization and analysis, Cytoscape v3.10.2 (Shannon et al. 2003) was utilized. This tool facilitated the construction, examination, and interpretation of subnetworks, allowing for the identification of key genes and their interactions.

Co-functional network analysis within the data systems layer provides a powerful approach to understanding how genes interact to mediate cold tolerance in soybean. By integrating SoyNet data with advanced visualization tools like Cytoscape, this layer refines the identification of key genes and functional modules, offering actionable insights for precision breeding and molecular studies aimed at improving cold tolerance in soybean.

Prioritization of key genes (data topology layer)

Data source

All topological metrics in this study were computed using the SoyNet database (Kim et al. 2017), which integrates multiple data types to construct biologically relevant and soybean-specific co-functional networks. The data pool included protein–protein interaction networks (PPIN) (experimental and predicted interactions), co-expression (RNA-seq and microarray datasets), pathway membership (KEGG and GO functional annotations), phylogenetic data (orthologous gene relationships), and curated literature (literature-curated interactions). These integrated data sources provided the edges (functional relationships) for the network and contributed to identifying tightly connected gene clusters. Co-expression data, pathway membership, and phylogenetic conservation were instrumental in detecting

biologically meaningful modules, while PPIN and curated literature strengthened the robustness of the network structure. This comprehensive integration ensures that topological metrics, such as degree centrality and clustering coefficient, accurately reflect the functional importance of genes in soybean-specific networks.

Topological analysis and prioritization

In the data topology layer, topological metrics were used to prioritize key CTgenes based on their roles within sub-networks. Two critical metrics—degree centrality and clustering coefficients—were applied to evaluate each gene's importance. Degree centrality measures the number of direct connections (edges) for a gene (node) in the network. Genes with high degree centrality are considered network hubs, playing essential roles in coordinating interactions and facilitating information flow across the network. Clustering coefficient quantifies the interconnectedness of a gene's neighbors, assessing the extent to which adjacent nodes are also interconnected. Genes with high clustering coefficients often belong to tightly connected modules or pathways that collectively mediate specific biological processes, such as cold tolerance. This approach combines biological relevance with network topology, enabling a comprehensive understanding of gene function.

Topological metrics help distinguish genes that are both biologically relevant and structurally central in the sub-network or gene modules, ensuring the selected genes contribute to key processes relevant to cold-tolerance and are integral to the system's overall robustness. Genes central to the network structure (i.e., with high scores in these metrics) are more likely to be functionally significant. These genes not only participate in specific biological pathways but also serve as hubs or key regulators that coordinate complex responses to cold stress. Hence, they were prioritized as key CTgenes for further experimental validation as they are hypothesized to play central roles in the regulatory networks underpinning cold tolerance in soybean. This layer ensures that both global (network-wide) and local (module-specific) roles of genes are considered in this systematic approach.

Validation for the key CTgenes (validation layer)

Statistical analysis

To assess the significance and performance of the identified key CTgenes, we compared their mean scores to those of other CTgenes, intermediate genes, and remaining genes under different cold treatment durations (1 h and 24 h). Statistical evaluation was conducted using an Analysis of

Variance (ANOVA) to determine whether significant difference ($p < 0.05$) existed among the four gene groups across short, middle, and combined networks. If significant differences were observed in ANOVA, a Tukey's Honestly Significant Difference (Tukey's HSD) test was applied for pairwise comparisons. This statistical framework ensures robust evaluation of gene prioritization, minimizing false positives while highlighting genes with the most prominent roles in cold-tolerance responses.

Molecular biology analysis

To validate the transcriptional activity of key CTgenes, soybean seeds of the Chiangmai 60 cultivar were surface-sterilized in 1% sodium hypochlorite, rinsed thoroughly with distilled water, and sown in sandy soil-filled pots (240 mm × 240 mm × 190 mm). Seedlings were cultivated under controlled conditions and watered regularly. Cold stress treatments followed a modified protocol from Yamasaki and Randall (2016), in which 10-day-old seedlings were exposed to 4 °C for 1, 3, 6, 12, and 24 h. Shoots and roots were harvested, immediately frozen in liquid nitrogen, and stored for transcript analysis. Each biological replicate included plant organs from at least four plants of a similar developmental stage, pooled from a minimum of two pots.

Total RNA was extracted using the TRIzol reagent following the manufacturer's protocol. RNA (5 µg) was reverse-transcribed into complementary DNA (cDNA) using Superscript™ III Reverse Transcriptase with oligo(dT)18 primers at 50 °C for 60 min. qRT-PCR was performed using the Luna® Universal qPCR Master Mix and gene-specific primers (Table S7). The actin gene was used as an internal control, and relative gene expression levels were calculated using the $2^{-\Delta\Delta CT}$ method, with untreated plant cDNA serving as the reference. All experiments were conducted in triplicate to ensure reproducibility. This molecular approach allows precise quantification of gene expression changes, enabling functional validation of key CTgenes involved in cold-tolerance responses.

Plant hormones analysis

To investigate hormonal changes associated with cold stress, soybean shoots were ground in liquid nitrogen, and 200 mg of the powder was extracted using a solvent mixture of methanol:water:formic acid (15:4:1, v/v) containing 10 ppm naphthalene acetic acid (NAA) as an internal standard. After vortexing for 5 min, the samples were incubated at -20 °C overnight. The supernatant was collected by centrifugation at $12,470 \times g$ for 15 min and subjected to solid-phase extraction (SPE) using Supel™ Select HLB cartridges. The eluate was prepared with

methanol:acetonitrile (50:50, v/v) containing 2% ammonium hydroxide and filtered through a 0.22 µm cellulose acetate syringe filter.

Hormone quantification was conducted using HPLC–MS/MS on a Shimadzu LCMS-8060 system with an Inertsil ODS-3 column (4.6 × 150 mm, 5 µm). The mobile phase consisted of buffer A (0.1% formic acid in water) and buffer B (0.1% formic acid in methanol) at a flow rate of 0.3 mL/min. A gradient program reduced buffer A from 65 to 5% over 35 min. Electrospray ionization (ESI) in both positive and negative modes with multiple reaction monitoring (MRM) was used to detect salicylic acid (SA), ABA, and other plant hormones. Among the detected hormones, SA and ABA were consistently quantified across samples. Hormone concentrations were calculated by comparing the peak areas of SA and ABA with standard curves (Fig. S1), using the formula:

Hormone concentration

$$= \frac{\text{Measured peak area} \times \text{NAA added concentration}}{\text{NAA measured concentration}}$$

Five biological replicates were analyzed for each sample to ensure statistical reliability. These results highlight the accumulation of stress-responsive hormones, such as SA and ABA, providing critical biochemical evidence of their roles in cold tolerance. The integration of this hormonal analysis with molecular validation offers a comprehensive understanding of the physiological and molecular mechanisms underlying soybean cold-tolerance responses.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00299-025-03643-2>.

Acknowledgements We are grateful to Prof. Dr. James R. Ketudat-Cairns for facilitating standard plant hormones.

Author contributions C.-F.K.: conceived and designed the project. C.-F.K., P.-H.K., and H.-Y.L. collected the data. S.B.: established the experimental data. H.-Y.L.: performed the data integration, analysis, and visualization. H.-Y.L., C.-F.K., and S.-B.: interpreted the data and the results. H.-Y.L. and C.-F.K.: drafted the manuscript. C.-F.K., and H.-Y.L.: revision of the manuscript. All authors read and approved the final manuscript.

Funding This research was supported by the 2-institution Co-Research Scholarship provided by Kasetsart University and National Chung Hsing University (No.00042023). This study was funded by Chung Cheng Agriculture Science and Social Welfare Foundation, and The Advanced Plant and Food Crop Biotechnology Center from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education in Taiwan. SB was supported by the Kasetsart University Research and Development Institute [FF(S-KU) 43.66] and partially supported by the Faculty of Science at Sriracha, Kasetsart University at Sriracha Campus.

Data availability The data supporting the results are available in this paper and its supplementary files.

Declarations

Conflict of interest The authors declare no conflicts of interest.

Ethical approval and consent to participate This article does not contain any studies with human participants or animals performed by any of the authors.

Consent for publication Not applicable. This article does not contain any studies with human participants or animals performed by any of the authors that require consent for publication.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

- Achard P, Gong F, Cheminant S, Alioua M, Hedden P, Genschik P (2008) The cold-inducible CBF1 factor-dependent signaling pathway modulates the accumulation of the growth-repressing DELLA proteins via its effect on gibberellin metabolism. *Plant Cell* 20:2117–2129
- Ackermann M, Strimmer K (2009) A general modular framework for gene set enrichment analysis. *BMC Bioinformatics* 10:47
- Afzal Z, Howton T, Sun Y, Mukhtar MS (2016) The roles of aquaporins in plant stress responses. *J Dev Biol* 4:9
- Badole S, Bodhankar S (2012) *Glycine max* (soybean) treatment for diabetes. *Bioact Food Dietary Intervent Diabetes* 77:77–82
- Ballouz S, Verleyen W, Gillis J (2015) Guidance for RNA-seq co-expression network construction and analysis: safety in numbers. *Bioinformatics* 31:2123–2130
- Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I (2001) Controlling the false discovery rate in behavior genetics research. *Behav Brain Res* 125:279–284
- Cheng L, Zhu M (2021) First-order correction of statistical significance for screening two-way epistatic interactions. *Methods Mol Biol* 2212:181–190
- Chinnusamy V, Zhu JK, Sunkar R (2010) Gene regulation during cold stress acclimation in plants. *Methods Mol Biol* 639:39–55
- Cirillo E, Parnell LD, Evelo CT (2017) A review of pathway-based analysis tools that visualize genetic variants. *Front Genet* 8:174
- Collard BC, Jahufer M, Brouwer J, Pang ECK (2005) An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. *Euphytica* 142:169–196
- Consortium TU (2024) UniProt: the universal protein knowledgebase in 2025. *Nucleic Acids Res* 53:D609–D617

- Csárdi G, Nepusz T, Müller K, Horvát S, Traag V, Zanini F, Noom D (2023) igraph for R: R interface of the igraph library for graph theory and network analysis. Zenodo
- Cuevas JC, López-Cobollo R, Alcázar R, Zarza X, Koncz C, Altabella T, Salinas J, Tiburcio AF, Ferrando A (2008) Putrescine is involved in Arabidopsis freezing tolerance and cold acclimation by regulating abscisic acid levels in response to low temperature. *Plant Physiol* 148:1094–1105
- De La Torre AR, Wilhite B, Puiu D, St. Clair JB, Crepeau MW, Salzberg SL, Langley CH, Allen B, Neale DB (2021) Dissecting the polygenic basis of cold adaptation using genome-wide association of traits and environmental data in Douglas-fir. *Genes* 12:110
- Ding F, Wang X, Li Z, Wang M (2022) Jasmonate positively regulates cold tolerance by promoting ABA biosynthesis in tomato. *Plants* 12:60
- Efron B, Tibshirani R (2002) Empirical bayes methods and false discovery rates for microarrays. *Genet Epidemiol* 23:70–86
- Goeman JJ, Buhlmann P (2007) Analyzing gene expression data in terms of gene sets: methodological issues. *Bioinformatics* 23:980–987
- Hesketh J, Myhre D, Willey C (1973) Temperature control of time intervals between vegetative and reproductive events in soybeans 1. *Crop Sci* 13:250–254
- Hu Y, Jiang L, Wang F, Yu D (2013) Jasmonate regulates the inducer of CBF expression—c-repeat binding factor/DRE binding factor1 cascade and freezing tolerance in Arabidopsis. *Plant Cell* 25:2907–2924
- Hwarari D, Guan Y, Ahmad B, Movahedi A, Min T, Hao Z, Lu Y, Chen J, Yang L (2022) ICE-CBF-COR signaling cascade and its regulation in plants responding to cold stress. *Int J Mol Sci* 23:1549
- Kao P-H, Baiya S, Lai Z-Y, Huang C-M, Jhan L-H, Lin C-J, Lai Y-S, Kao C-F (2022) An advanced systems biology framework of feature engineering for cold tolerance genes discovery from integrated omics and non-omics data in soybean. *Front Plant Sci* 13:1019709
- Kim JC, Lee SH, Cheong YH, Yoo CM, Lee SI, Chun HJ, Yun DJ, Hong JC, Lee SY, Lim CO (2001) A novel cold-inducible zinc finger protein from soybean, SCOF-1, enhances cold tolerance in transgenic plants. *Plant J* 25:247–259
- Kim MY, Van K, Kang YJ, Kim KH, Lee S-H (2012) Tracing soybean domestication history: from nucleotide to genome. *Breed Sci* 61:445–452
- Kim E, Hwang S, Lee I (2017) SoyNet: a database of co-functional networks for soybean *Glycine max*. *Nucleic Acids Res* 45:D1082–D1089
- Kokubun M (2011) Physiological mechanisms regulating flower abortion in soybean. In: *Soybean-biochemistry, chemistry and physiology*. Citeseer
- Lee YG, Jeong N, Kim JH, Lee K, Kim KH, Pirani A, Ha BK, Kang ST, Park BS, Moon JK (2015) Development, validation and genetic analysis of a large soybean SNP genotyping array. *Plant J* 81:625–636
- Leff B, Ramankutty N, Foley JA (2004) Geographic distribution of major crops across the world. *Glob Biogeochem Cycles*.
- Lissarre M, Ohta M, Sato A, Miura K (2010) Cold-responsive gene regulation during cold acclimation in plants. *Plant Signal Behav* 5:948–952
- Mariola Staniak KC, Stepień-Warda A, Kocira A, Przybyś M (2021) Cold stress during flowering alters plant structure, yield and seed quality of different soybean genotypes. *Agron J* 11:2059
- Mirina A, Atzmon G, Ye K, Bergman A (2012) Gene size matters. *PLoS ONE* 7:e49093
- Ohnishi S, Miyoshi T, Shirai S (2010) Low temperature stress at different flower developmental stages affects pollen development, pollination, and pod set in soybean. *Environ Exp Bot* 69:56–62
- Priyanatha C, Torkamaneh D, Rajcan I (2022) Genome-wide association study of soybean germplasm derived from Canadian× Chinese crosses to mine for novel alleles to improve seed yield and seed quality traits. *Front Plant Sci* 13:866300
- Qian Z, He L, Li F (2024) Understanding cold stress response mechanisms in plants: an overview. *Front Plant Sci* 15:1443317
- Rahmatallah Y, Emmert-Streib F, Glazko G (2012) Gene set analysis for self-contained tests: complex null and specific alternative hypotheses. *Bioinformatics* 28:3073–3080
- Ray DK, West PC, Clark M, Gerber JS, Prishchepov AV, Chatterjee S (2019) Climate change has likely already affected global food production. *PLoS ONE* 14:e0217148
- Raza A, Charagh S, Najafi-Kakavand S, Abbas S, Shoaib Y, Anwar S, Sharifi S, Lu G, Siddique KH (2023) Role of phytohormones in regulating cold stress tolerance: physiological and molecular approaches for developing cold-smart crop plants. *Plant Stress* 8:100152
- Raza A, Bashir S, Khare T, Karikari B, Copeland RG, Jamla M, Abbas S, Charagh S, Nayak SN, Djalovic I (2024) Temperature-smart plants: a new horizon with omics-driven plant breeding. *Physiol Plant* 176:e14188
- Raza A, Li Y, Prakash CS, Hu Z (2025) Panomics to manage combined abiotic stresses in plants. *Trends Plant Sci*.
- Ren X, Hu Q, Liu S, Wang J, Miecznikowski JC (2017) Gene set analysis controlling for length bias in RNA-seq experiments. *BioData Min* 10:5
- Roychowdhury R, Das SP, Gupta A, Parihar P, Chandrasekhar K, Sarker U, Kumar A, Ramrao DP, Sudhakar C (2023) Multi-omics pipeline and omics-integration approach to decipher plant's abiotic stress tolerance responses. *Genes* 14:1281
- Ruelland E, Vaultier M-N, Zachowski A, Hurry V (2009) Chapter 2 cold signalling and cold acclimation in plants. *Adv Bot Res* 49:35–150
- Sadura I, Janeczko A (2024) Are heat shock proteins important in low-temperature-stressed plants? A minireview. *Agronomy* 14:1296
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–2504
- Song W, Shao H, Zheng A, Zhao L, Xu Y (2023) Advances in roles of salicylic acid in plant tolerance responses to biotic and abiotic stresses. *Plants* 12:3475
- Staniak M, Czopek K, Stepień-Warda A, Kocira A, Przybyś M (2021) Cold stress during flowering alters plant structure, yield and seed quality of different soybean genotypes. *Agronomy* 11:2059
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP (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
- Survarachakan S, Prasad PJR, Naseem R, Pérez de Frutos J, Kumar RP, Langø T, Alaya Cheikh F, Elle OJ, Lindseth F (2022) Deep learning for image-based liver analysis — A comprehensive review focusing on malignant lesions. *Artif Intell Med* 130:102331
- Szabados L, Savaure A (2010) Proline: a multifunctional amino acid. *Trends Plant Sci* 15:89–97
- Thomashow MF (2010) Molecular basis of plant cold acclimation: insights gained from studying the CBF cold response pathway. *Plant Physiol* 154:571–577
- Tian F, Yang D-C, Meng Y-Q, Jin J, Gao G (2020) Plantregmap: charting functional regulatory maps in plants. *Nucleic Acids Res* 48:D1104–D1113
- Valdés-Pérez RE (1999) Discovery tools for science apps. *Commun ACM* 42:37–41
- Visscher PM, Brown MA, McCarthy MI, Yang J (2012) Five years of GWAS discovery. *Am J Hum Genet* 90:7–24
- Wang M, Fan X, Ding F (2023) Jasmonate: a hormone of primary importance for temperature stress response in plants. *Plants* 12:4080

- Wang Q, Qi C, Wang L, Li M, Niu Y, Muhammad N, Liu M, Liu Z, Wang L (2025) ZjMAPKK4 interacted with ZjNAC78 regulates cold tolerance response in jujube. *Plant, Cell Environ.*
- Wasternack C, Hause B (2013) Jasmonates: biosynthesis, perception, signal transduction and action in plant stress response, growth and development. An update to the 2007 review in *Annals of Botany*. *Ann Bot* 111:1021–1058
- Wei W, Wu X, Blahut-Beatty L, Simmonds DH, Clough SJ (2022) Transcriptome profiling reveals molecular players in early soybean-*sclerotinia sclerotiorum* interaction. *Phytopathology*® 112:1739–1752
- Yamasaki Y, Randall SK (2016) Functionality of soybean CBF/DREB1 transcription factors. *Plant Sci* 246:80–90
- Yu W, Yu W, Yang Y, Lü Y (2021) Exploring the key genes and identification of potential diagnosis biomarkers in Alzheimer's disease using bioinformatics analysis. *Front Aging Neurosci* 13:602781
- Zeng Z, Zhang S, Li W, Chen B, Li W (2022) Gene-coexpression network analysis identifies specific modules and hub genes related to cold stress in rice. *BMC Genomics* 23:251
- Zhang J, Song Q, Cregan PB, Nelson RL, Wang X, Wu J, Jiang G-L (2015) Genome-wide association study for flowering time, maturity dates and plant height in early maturing soybean (*Glycine max*) germplasm. *BMC Genomics* 16:1–11
- Zhao C, Liu X, He J, Xie Y, Xu Y, Ma F, Guan Q (2021) Apple TIME FOR COFFEE contributes to freezing tolerance by promoting unsaturation of fatty acids. *Plant Sci* 302:110695

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.